

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

cleic search, using sw model

April 13, 2004, 18:17:02 ; Search time 1853.5 Seconds  
(without alignments)  
1240.565 Million cell updates/sec

US-10-090-326-25

1 acttacnagccagcancan.....ggagnacaacagccancagt 77

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

27513289 seqs, 14931090276 residues

hits satisfying chosen parameters: 55026578

length: 0

length: 2000000000

: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

EST:\*

1: em estba:\*

2: em esthum:\*

3: em estin:\*

4: em estma:\*

5: em estov:\*

6: em estpl:\*

7: em estro:\*

8: em htc:\*

9: gb est1:\*

10: gb est2:\*

11: gb htc:\*

12: gb est3:\*

13: gb est4:\*

14: gb est5:\*

15: em estfun:\*

16: em estom:\*

17: em gss hum:\*

18: em gss inv:\*

19: em gss pln:\*

20: em gss vrt:\*

21: em gss fun:\*

22: em gss nam:\*

23: em gss mus:\*

24: em gss pro:\*

25: em gss rod:\*

26: em gss phg:\*

27: em gss vrl:\*

28: gb gss1:\*

29: gb gss2:\*

is the number of results predicted by chance to have a  
iter than or equal to the score of the result being printed,  
rived by analysis of the total score distribution.

# SUMMARIES

Query	Match	Length	DB	ID	Description
49.1	939	12	BG764646	BG764646	602736726
39.2	522	28	AQ585695	AQ585695	RPCI-11-4
39.0	294	13	BQ678960	BQ678960	AGENCOURT
39.0	390	14	N28965	N28965	yx39d02.r1

5	30	39.0	674	13	BX102971	BX102971
6	30	39.0	764	12	BI859609	BI859609
7	30	39.0	908	13	BQ424713	BQ424713
8	30	39.0	926	13	BQ424642	BQ424642
9	30	39.0	943	13	BQ677071	BQ677071
10	30	39.0	965	12	BG576440	BG576440
11	29.8	38.7	1083	12	BM463960	BM463960
12	29.2	37.9	1155	13	BX340723	BX340723
13	28.8	37.4	461	28	AZ095479	AZ095479
14	28.8	37.4	651	28	AZ661932	AZ661932
15	28.6	37.1	935	10	BF237982	BF237982
16	28.2	36.6	400	13	BX749335	BX749335
17	28	36.4	617	28	AZ658637	AZ658637
18	28	36.4	625	29	CE510405	CE510405
19	28	36.4	939	13	BU465668	BU465668
20	27.6	35.8	544	28	AZ135044	AZ135044
21	27.6	35.8	781	28	AQ860588	AQ860588
22	27.4	35.6	221	10	BS588106	BS588106
23	27.4	35.6	301	9	AU183092	AU183092
24	27.4	35.6	304	29	CE670554	CE670554
25	27.4	35.6	330	9	AU182930	AU182930
26	27.4	35.6	447	9	AU183801	AU183801
27	27.4	35.6	451	9	AU223051	AU223051
28	27.4	35.6	453	9	AU222931	AU222931
29	27.4	35.6	454	9	AU183799	AU183799
30	27.4	35.6	470	9	AU172933	AU172933
31	27.4	35.6	482	12	BM487714	BM487714
32	27.4	35.6	687	14	CB685142	CB685142
33	27.4	35.6	782	29	CB474611	CB474611
34	27.4	35.6	984	13	BU190117	BU190117
35	27.4	35.6	1068	12	BM557397	BM557397
36	27.2	35.3	462	12	BM595860	BM595860
37	27.2	35.3	465	12	BM644507	BM644507
38	27.2	35.3	481	12	BM648778	BM648778
39	27.2	35.3	482	11	CNS08F9W	CNS08F9W
40	27.2	35.3	487	12	BM643718	BM643718
41	27.2	35.3	503	12	BM655292	BM655292
42	27.2	35.3	507	12	BM622186	BM622186
43	27.2	35.3	510	12	BM620575	BM620575
44	27.2	35.3	512	12	BM587290	BM587290
45	27.2	35.3	541	12	BM644366	BM644366

## ALIGNMENTS

RESULT 1	BG764646	BG764646	602736726	Fl NIH_MGC_49	Homo sapiens	cDNA clone IMAGE:48	linear	EST
LOCUS	BG764646							
DEFINITION	BG764646							
ACCESSION	BG764646							
VERSION	BG764646.1							
KEYWORDS	EST.							
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							
REFERENCE	1	(bases 1 to 939)						
AUTHORS	NIH-MGC	http://mgc.nci.nih.gov/						
TITLE	National Institutes of Health, Mammalian Gene Collection							
JOURNAL	Unpublished (1999)							
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue Procurement: ATCC/DCTD/DTF cDNA Library Preparation: Ling Hong/Rubin Laboratory DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLM1721 row: d column: 07 High quality sequence stop: 885.							

```

Location/Qualifiers
1. .939
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4861782"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_49"
/notes="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. ["

Query Match 39.2%; Score 30.2; DB 28; Length 522;
Best Local Similarity 62.0%; Pred. No. 20;
Matches 44; Conservative 0; Mismatches 27; Indels 0

ORIGIN
QY 5 ACNACGCCAGCAGCATCTAGCATCATCCTCTGCATGGTCAGGTCATTGGT
Db 485 ACACGCCAGGTCAGTATCTCGTTGTCATATCTTTGCACAGCAGTTTC
QY 65 AACAGCCANCA 75
Db 425 AACAGCCACCA 415

RESULT 3
BQ678960
LOCUS
DEFINITION BQ678960 294 bp mRNA linear EST
5', mRNA sequence.
ACCESSION BQ678960
VERSION BQ678960.1 GI:21791639
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hominid;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eul
1 (bases 1 to 294)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection
Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LI
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information f
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2423 row: 1 column: 24
High quality sequence stop: 293.
Location/Qualifiers
1. .294
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6261983"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_112"
/notes="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dt priming. Directic
into EcoRI/XhoI sites using the following 5' a
GGCACAG(G). Library constructed by Ling Hong
laboratory of Gerald M. Rubin (University of
Berkeley) using ZAP-cDNA synthesis kit (Strata
Superscript II RT (Life Technologies). Note: this
NIH_MGC Library. ["

Query Match 39.0%; Score 30; DB 13; Length 294;
Best Local Similarity 67.9%; Pred. No. 18;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

ORIGIN
QY 1 ACTTACNAGCCAGCAGCATCTAGCATCATCCTCTGCATGGTCAGGTCATTGGT

```

ACTTACTAGCCAGCCATCATCTCTCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 268

AGNACACAGCCANCACT 77  
|||||  
AGTACACAGCCATCAGT 286

8965 390 bp mRNA linear EST 04-JAN-1996  
39002.r1 Soares melanocyte 2NBHM Homo sapiens cDNA clone  
AGE:264099 5' similar to gb:M27160\_rnal TYROSINASE PRECURSOR  
UNAN); mRNA sequence.  
8965  
8965.1 GI:1147201  
T.  
mo sapiens (human)  
no sapiens  
karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
mmlia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
(bases 1 to 390)  
llier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
lian,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
rsos,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,  
evaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and  
lson,R.  
e WashU-Merck EST Project  
published (1995)  
nact: Wilson RK  
shington University School of Medicine  
44 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
1: 314 286 1800  
x: 314 286 1810  
ail: est@watson.wustl.edu  
gh quality sequence stops: 296  
urce: IMAGE Consortium, LNL  
is clone is available royalty-free through LNL; contact the  
AGE Consortium (info@image.llnl.gov) for further information.  
g primer: T7  
gh quality sequence stop: 296.

Location/Qualifiers  
1..390  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:387374.1"  
/db\_xref="taxon:9606"  
/clone="IMAGE:264099"  
/sex="Male"  
/tissue\_type="melanocyte"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares melanocyte 2NBHM"  
/note="Vector: pT73D (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5'  
TGTTCACCAATCTGAAGTGGAGCGCGCCGAGTTTCTTTTCTTTT 3'],  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pT73 vector  
(Pharmacia). Library constructed by Bento Soares and  
M.Fatima Bonaldo. RNA from normal foreskin melanocytes  
(FS374) was kindly provided by Dr. Anthony P. Albino."

39.0%; Score 30; DB 14; Length 390;  
imilarity 67.9%; Pred. No. 21;  
; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

ACTTACAGCCAGCCANCACTTACATCATCTCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 158  
ACTTACTAGCCAGCCATCATCTCTCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 158

AGNACACAGCCANCACT 77  
|||||

Db 159 AGTACACAGCCATCAGT 176

RESULT 5  
BX102971  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

674 bp mRNA linear E;  
BX102971 Soares melanocyte 2NBHM Homo sapiens cDNA cl;  
IMAGE:264099; mRNA sequence.  
BX102971  
BX102971.1 GI:27832125  
EST.  
Homo sapiens (human)  
Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; E;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H  
1 (bases 1 to 674)  
Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E.,  
Radelof,U., Schneider,D. and Korn,B.  
Human Unigeneset - RZPD3  
Unpublished (2003)  
Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung C  
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
RZPD; IMAGE:998B04578  
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;  
Human Unigeneset - RZPD3 (RZPDLIB No.972)  
http://www.rzpd.de/CloneCards/cgi-  
bin/showLib.pl.cgi/?response?libNo=972 Contact: Ina Rol  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung C  
Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
www.rzpd.de

This clone is available royalty-free from RZPD;  
contact RZPD (clone@rzpd.de) for further information.  
M13r, Primer sequence: TTTACACAGGAAACAGCTAGAC.

FEATURES  
source

Location/Qualifiers  
1..674  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:998B04578 ; IMAGE:264099"  
/sex="Male"  
/tissue\_type="melanocyte"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares melanocyte 2NBHM"  
/note="Vector: pT73D (Pharmacia) with a modi  
polylinker; Site 1: Not I; Site 2: Eco RI; 1s  
was primed with a Not I - oligo(dT) primer [5  
TGTTCACCAATCTGAAGTGGAGCGCGCCGAGTTTCTTTTCTTTT  
double-stranded cDNA was size selected, ligat  
adapters (Pharmacia), digested with Not I and  
the Not I and Eco RI sites of a modified pT77  
(Pharmacia). Library constructed by Bento Soa  
M.Fatima Bonaldo. RNA from normal foreskin me  
(FS374) was kindly provided by Dr. Anthony P.

ORIGIN

Query Match 39.0%; Score 30; DB 13; Length 674;  
Best Local Similarity 67.9%; Pred. No. 27;  
Matches 53; Conservative 0; Mismatches 24; Indels 1

QY 1 ACTTACAGCCAGCCANCACTTACATCATCTCTCTTGGCAGATTGTCTGTAGCGTCA  
|||||

Db 127 ACTTACTAGCCAGCCATCATCTCTCTCTTGGCAGATTGTCTGTAGCGG

QY 60 AGNACACAGCCANCACT 77  
|||||

Db 187 AGTACACAGCCATCAGT 204

RESULT 6  
BI859609







```

/db_xref="taxon:9606"
/clone="CS0DJ005Y008"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-I
primer. Five prime end enriched, double-strand
digested with Not I and cloned into the Not I
sites of the pCMVSPORT 6 vector. Library was n
ORIGIN
Query Match 37.9%; Score 29.2; DB 13; Length 1155;
Best Local Similarity 52.8%; Pred. No. 63;
Matches 38; Conservative 8; Mismatches 26; Indels 0;
QY 5 A NCACGCCAGCANCATTTTCAGCATCATCTCTGTCATGTCAGTCATTTGGGA
Db 1052 ACCAVACACGACGACCTKTCTGGCGCATTTTTCASAAATYYCATCTCAGTGGG
QY 65 AACAGCCANCAG 76
Db 992 AGSAGRSAGCAR 981
RESULT 13
AZ095479/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1..461
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="Rpci-23-474A7"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6;
ECORI; Site 2: EORI; Female C57BL/6J mouse ki
brain genomic DNA was isolated and partially

```



ACTTACNAGCCAGCCAGCANCATCTAGCATCATCTCTGCAATGCTCAGGTCAATTTGGAGGA 60  
 |||||  
 ACGTCCAGGCTGTTCCACTCCAGCTGAGGCTCTGAATGACTTCTTCTTTGGAGGA 561  
 |||||  
 : 61  
 : 560

749935 400 bp mRNA linear EST 09-DEC-2003  
 749935 XGC-gastrula Silurana tropicalis cDNA clone Tgas073m07 3',  
 5A sequence.  
 749935  
 749935.1 GI:39640003

..urana tropicalis (western clawed frog)  
 ..urana tropicalis  
 ..aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 ..hibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 ..opodinae; Silurana.  
 (bases 1 to 400)  
 ming,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.  
 iger Xenopus tropicalis EST project 2001 (11\_2003)  
 ublished (2003)  
 tact: Croning MDR  
 iger Institute  
 xton, Cambridgeshire, CB10 1SA, UK  
 il: trop@sanger.ac.uk  
 iger Xenopus tropicalis EST project 2001  
 PICALIS SEQUENCE ID: Tgas073m07.q1kT7  
 uencing primer: T7  
 s sequence is from a Xenopus Gene Collection (XGC) library  
 irected by Aaron M. Zorn.  
 1A was oligo dT primed from Sug of poly A+ RNA from stages 10-13  
 trulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with  
 RI at the 5' end and NotI at the 3' end.  
 tor: pCS107; Site 1: EcoRI; Site 2: NotI  
 it: Escherichia coli XL1-blue.  
 Location/Qualifiers  
 1..400  
 /organism="Silurana tropicalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8364"  
 /clone="Tgas073m07"  
 /dev\_stage="gastrula (stages 10.5-12 mixed)"  
 /lab\_host="Escherichia coli XL1-blue"  
 /clone\_lib="XGC-gastrula"  
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA  
 was oligo dT primed from Sug of poly A+ RNA from stages  
 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated  
 into pCS107 with EcoRI at the 5' end and NotI at the 3'  
 end."

36.6%; Score 28.2; DB 13; Length 400;  
 milarity 65.0%; Pred.No. 84;  
 Conservative 0; Mismatches 21; Indels 0; Gaps 0;

ACKNAGCCAGCCAGCANCATCTAGCATCATCTCTGCAATGCTCAGGTCAATTTGGAGGAGNA 63  
 |||||  
 ACCCATAGACCACTTACGACCGTGCCTTATGCTCAGATCATTTGTGTGGA 96  
 |||||

58637 617 bp DNA linear GSS 14-DEC-2000  
 535105R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
 ne UUGC1M0535105 R, genomic survey sequence.  
 58637  
 58637.1 GI:11795783

KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; M  
 1 (bases 1 to 617)  
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., H  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen  
 Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von  
 Niederhausen,A. and Wright,D.,Weiss,R.  
 Mouse whole genome scaffolding with paired end reads f  
 plasmid inserts  
 Unpublished (2000)  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 338, Biomedical Polymers Research Bldg., 20 S. 203E  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0535 row: I column: 05  
 Seq primer: CACACGGAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 617.  
 Location/Qualifiers  
 1..617  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0535105"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resist  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /note="Vector: PWD42nv; Purified genomic DNA f  
 musculus C57BL/6J (male) was obtained from th  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares  
 was hydrodynamically sheared by repeated pass  
 0.005 inch orifice at constant velocity. The  
 was blunt end-repaired with T4 DNA polymerase  
 polynucleotide kinase. Adaptor oligonucleotide  
 ligated to the blunt ends in high molar exce  
 adaptor DNA was purified and size-selected f  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from  
 of PWD42 (G14732114|G14732114|G14732114), a copy-nt  
 inducible derivative of plasmid R1. The vector  
 with adaptors complementary to the insert adap  
 purified. The sheared, adaptor mouse DNA was  
 adaptor vector DNA, and transformed into  
 chemically-competent E. coli XL10-Gold (Strata  
 and selected for ampicillin resistance."

ORIGIN  
 Query Match 36.4%; Score 28; DB 28; Length 617;  
 Best Local Similarity 60.6%; Pred. No. 1.2e+02;  
 Matches 43; Conservative 0; Mismatches 28; Indels 0;  
 Qy 1 ACTTACNAGCCAGCANCATCTAGCATCATCTCTGCAATGCTCAGGTCAATTTGGAGGAGNA 63  
 Db 389 ACTTGTCTCTCAGCAGCGTGAAGCTTCATTTCGAGAGCGGCGGAGGTCAAG  
 |||||  
 Qy 61 GNACACAGCC 71  
 Db 449 GTACAGAGGCC 459  
 |||||  
 RESULT 18  
 CE510405/c  
 LOCUS CE510405 625 bp DNA linear GSS



## Location/Qualifiers

```

1. .544
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice BAC Library (EcoRI)"
/note="Vector: pBACindigo; Site 1: EcoRI; Site 2: EcoRI; Rice is the most important food crop in the world. Half of the world population, especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa. Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 Kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center (www.genome.clemson.edu)."

```

```

35.8%; Score 27.6; DB 28; Length 544;
similarity 65.0%; Pred. No. 1.5e+02;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```

```

TTACNCGCCGACGACATTCATGACATCATCTCTGCGATGTCAGTCTATTTGGAGGAG 61
|||||
TTGGAGTGCACGTCTTATTAACGTCATCATCTCTGATGATCAGGGCACATGGTGCAG 208

```

```

60588      781 bp      DNA      linear      GSS 03-NOV-1999
b0015K14r CUGI Rice BAC Library (EcoRI) Oryza sativa (japonica
tivar-group) genomic clone nb0015K14r, genomic survey
uence.

```

```

60588      60588.1      GI:6211045

```

```

za sativa (japonica cultivar-group)
za sativa (japonica cultivar-group)
aryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
matophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
hartodeae; Oryzaceae; Oryza.
(bases 1 to 781)
g.R.A. and Dean,R.A.
AC End Sequencing Framework to Sequence the Rice Genome
ublished (1998)
tact: Wing RA
mson University Genomics Institute
Jordan Hall, Clemson, SC 29634, USA
: 864 656 7288
: 864 656 4293
til: rwing@clemson.edu
l primer: GGAAACAGCTATGACCATG
ss: BAC ends
th quality sequence start: 37
th quality sequence stop: 408.
Location/Qualifiers

```

## source

```

1. .781
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="nb0015K14r"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice BAC Library (EcoRI)"
/note="Vector: pBACindigo; Site 1: EcoRI; Site 2: EcoRI; Rice is the most important food crop in the world. Half of the world population, especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa. Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 Kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center (www.genome.clemson.edu)."

```

## ORIGIN

```

Query Match      35.8%; Score 27.6; DB 28; Length 781;
Best Local Similarity 65.0%; Pred. No. 1.8e+02;
Matches 39; Conservative 0; Mismatches 21; Indels 0;

```

```

QY      2      CTTCACGCCGACGACATTCATGACATCATCTCTGCGATGTCAGTCTATTT
Db      326      CTGGAGTGCACGTCTTATTAACGTCATCATCTCTGATGATCAGGGCACAT

```

## RESULT 22

```

BB588106/c
LOCUS
DEFINITION
musculus cDNA clone AC30005102 5', mRNA sequence.
BB588106
BB588106.1      GI:11484650
EST.
Mus musculus (house mouse)

```

```

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mu
1 (bases 1 to 221)
Aizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T.,
Carninci, P., Hanagaki, T., Hayatsu, N., Hiraoka, T., Hiroz
Hodoyama, Y., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., K
Kojima, Y., Konno, H., Kusakabe, M., Matsuyama, T., Miyazaki
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Okazaki
Okido, T., Owa, C., Sakai, K., Sakai, K., Sasaki, D., Sato, K
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogab
Suzuki, H., Tagawa, A., Takahashi, F., Tanaka, T., Toya, T.,
Watahiki, A., Yamamura, T., Yasunishi, A., Yoshida, K., Yos
Muranatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Aizawa, K. et al. 2000)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-
Tel: 81-45-503-9222

```

```

Mus musculus (house mouse)

```

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

```

: 81-45-503-9216
ii: genome-resgsc.riken.go.jp,
: http://genome-gsc.riken.go.jp/
ninci,P., Nishiyama,Y., Westover,A., Itch,M., Nagaoka,S.,
aki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
rmotabilization and thermoactivation of thermolabile enzymes by
halose and its application for the synthesis of full length
A. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Oh,M., Kitsuai,I., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
aru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
zaki,Y. and Hayashizaki,Y.
tomated filtration-based high-throughput plasmid preparation
tem. Genome Res. 9 (5), 463-470 (1999)
rhinci,P. and Hayashizaki,Y.
gh-efficiency full-length cDNA cloning. Methods Enzymol. 303,
44 (1999)
ease visit our web site (http://genome.rtc.riken.go.jp) for
their details.
Location/Qualifiers
1..221
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A030005102"
/tissue_type="skin"
/dev_stage="6 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 6 days neonate
skin"
notes="Site 1: Sall; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGCTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 185.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCGAGTTAATTAATTAATCCGCCGCCGCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC 1."
nilarity 35.6%; Score 27.4; DB 10; Length 221;
Conservative 0; Mismatches 29; Indels 0; Gaps 0;
TTATCNCAGCCAGCAGCATTTAGCATCATCTCTGCGATGTCAGGTCTGGAGGA 60
|||
TTAATTCGAGATATACACTTTAGCATGATTTTATCATGTCGAAGCAGCATTTGGAGT 147
|||
NACAACAGCCA 72
|||
GACTAGAGACA 135

83092 301 bp mRNA linear EST 03-APR-2002
83092 Rice panicle (between 3cm to 10cm) Oryza sativa (japonica
tivar-group) cDNA clone E41124, mRNA sequence.
83092
83092.1 GI:13896756
:
za sativa (japonica cultivar-group)
za sativa (japonica cultivar-group)
aryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
rnatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 301)
Sasaki,T. and Yamamoto,K.
Rice cDNA from panicle (between 3cm to 10cm) (2001)
Unpublished (2001)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba,
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affr
PROJECT = 'RGP'
FEATURES
Location/Qualifiers
source
1..301
/organism="Oryza sativa (japonica cultivar-grc
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="E41124"
/tissue_type="panicle"
/dev_stage="panicle (between 3cm to 10cm)"
/clone_lib="Rice panicle (between 3cm to 10cm)"
ORIGIN
Query Match 35.6%; Score 27.4; DB 9; Length 301;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 40; Conservative 0; Mismatches 24; Indels 0;
QY 12 CCAGCANCATTCCTAGCATCATCTCTGCGATGTCAGGTCTTTGGAGGAGNAC
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 CCACACATAGTATCATCAGAGTTTCTGCTAGTAGTATCTGACACAGAAC
QY 72 ANCA 75
|||
Db 181 AACA 178

RESULT 24
CE670554 304 bp DNA linear GSS
tigr-gss-dog-17000313750722 Dog Library Canis familiaris
genomic survey sequence.
CE670554
ACCESSION CE670554.1 GI:36989554
VERSION
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Can
1 (bases 1 to 304)
Kirkness,E.F., Barna,V., Halpern,A.L., Levy,S., Remingt
Rusch,D.B., Delcher,A.L., Pop,M., Wang,M., Fraser,C.M.,
Venter,J.C.
The dog genome: survey sequencing and comparative analy
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical C
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
FEATURES
Location/Qualifiers
source
1..304
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"

```

/note="Site 1: BstXI; Libraries were prepared from peripheral blood"

35.6%; Score 27.4; DB 29; Length 304;  
milarity 73.9%; Pred. No. 1.4e+02;  
Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
TCTAGCATCATCTCGCAGTGGTCAGGTCATTTGGAGGAGNACAA 66  
|||||  
TATGAATCATCTTCCATGGTGAAGCTTCTTTAGAGAGACAA 243  
|||||

82930 330 bp mRNA linear EST 03-APR-2002  
82930 Rice panicle (between 3cm to 10cm) Oryza sativa (japonica cultivar-group) cDNA clone E40377, mRNA sequence.  
82930  
82930.1 GI:13896594  
za sativa (japonica cultivar-group)  
za sativa (japonica cultivar-group)  
aryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
matophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
hartoideae; Oryzeae; Oryza.  
(bases 1 to 330)  
aki, T. and Yamamoto, K.  
e cDNA from panicle (between 3cm to 10cm) (2001)  
ublished (2001)  
tact: Takuji Sasaki  
e National Institute of Agrobiological Resources  
e Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
;-8602, Japan  
:: 81-298-38-7441  
:: 81-298-38-7468  
il: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
NJECT = 'RGP'.

Location/Qualifiers  
1. .330  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="E40377"  
/tissue\_type="panicle"  
/dev\_stage="panicle (between 3cm to 10cm)"  
/clone\_lib="Rice panicle (between 3cm to 10cm)"  
35.6%; Score 27.4; DB 9; Length 330;  
milarity 62.5%; Pred. No. 1.4e+02;  
Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
TCAGCANCATCTAGCATCTCTGCAGTGGTCAGGTCATTTGGAGGAGNACAGCC 71  
|||||  
CCACCATAGTATCATCAGAGTTTTTCTGTCATAGGTAATCTGAACAGACACAGGC 211  
|||||

ANCA 75

|||||

VACA 207

183801 447 bp mRNA linear EST 03-APR-2002  
183801 Rice panicle (between 3cm to 10cm) Oryza sativa (japonica cultivar-group) cDNA clone E40741, mRNA sequence.  
183801  
183801.1 GI:14191590  
za sativa (japonica cultivar-group)  
za sativa (japonica cultivar-group)

# REFERENCE

AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba,  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp  
PROJECT = 'RGP'.

## FEATURES

source  
Location/Qualifiers  
1. .447  
/organism="Oryza sativa (japonica cultivar-grc  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="E40741"  
/tissue\_type="panicle"  
/dev\_stage="panicle (between 3cm to 10cm)"  
/clone\_lib="Rice panicle (between 3cm to 10cm)"

## ORIGIN

Query Match 35.6%; Score 27.4; DB 9; Length 447;  
Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
Matches 40; Conservative 0; Mismatches 24; Indels 0;  
Qy 12 CCAGCANCATCTAGCATCATCTCTGCAGTGGTCAGGTCATTTGGAGGAGNAC  
|||  
Db 388 CCACCATAGTATCATCAGAGTTTTTCTGTCATAGGTAATCTGAACAGAC  
|||  
Qy 72 ANCA 75  
Db 328 VACA 325

## RESULT 27

AU223051/c  
LOCUS AU223051 Rice shoot Oryza sativa (japonica cultivar-grc  
DEFINITION clone S3117, mRNA sequence.  
ACCESSION AU223051  
VERSION AU223051.1 GI:15008663  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tr  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poac  
Ehrhartoideae; Oryzeae; Oryza.

## REFERENCE

AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba,  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp  
PROJECT = 'RGP'.

## FEATURES

source  
Location/Qualifiers  
1. .451  
/organism="Oryza sativa (japonica cultivar-grc  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="S3117"  
/dev\_stage="Etiolated shoot (8 days old)"



```

/clone_lib="Rice shoot"
/note="Etiolated shoot (8 days old)"

35.6%; Score 27.4; DB 9; Length 451;
imilarity 62.5%; Pred. No. 1.6e+02;
; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

CCAGCANCATTTAGCATCATCTCTGCATGTCAGGTCATTGGAGGAGNACAACAGCC 71
|||||
CCAAACCACATAGTATCATCAGAGTTTTTCTGCATAGGTAATCTGAACAGAGNACAACAGCC 185
|||||

ANCA 75
|||
AAC 181

222931 453 bp mRNA linear EST 03-APR-2002
222931 Rice shoot Oryza sativa (japonica cultivar-group) cDNA
one S1940, mRNA sequence.
222931
222931.1 GI:15008543
T.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Karyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaeae; Oryza.
(bases 1 to 453)
Sasaki, T. and Yamamoto, K.
cDNA from etiolated shoot (2001)
published (2001)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
5-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
PROJECT = 'RGP'.
Location/Qualifiers
1..453
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="S1940"
/dev_stage="Etiolated shoot (8 days old)"
/clone_lib="Rice shoot"
/note="Etiolated shoot (8 days old)"

35.6%; Score 27.4; DB 9; Length 453;
imilarity 62.5%; Pred. No. 1.6e+02;
; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

CCAGCANCATTTAGCATCATCTCTGCATGTCAGGTCATTGGAGGAGNACAACAGCC 71
|||||
CCAAACCACATAGTATCATCAGAGTTTTTCTGCATAGGTAATCTGAACAGAGNACAACAGCC 342
|||||

ANCA 75
|||
AAC 338

183799 454 bp mRNA linear EST 03-APR-2002
183799 Rice panicle (between 3cm to 10cm) Oryza sativa (japonica
cultivar-group) cDNA clone E40701, mRNA sequence.
183799

```

```

VERSION AUI83799.1 GI:14191588
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; T
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poac
Ehrhartoideae; Oryzaeae; Oryza.
1 (bases 1 to 454)
Sasaki, T. and Yamamoto, K.
Rice cDNA from panicle (between 3cm to 10cm) (2001)
Unpublished (2001)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.aff
PROJECT = 'RGP'.
Location/Qualifiers
1..454
/organism="Oryza sativa (japonica cultivar-gr
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="E40701"
/tissue_type="panicle"
/dev_stage="panicle (between 3cm to 10cm)"
/clone_lib="Rice panicle (between 3cm to 10cm)

ORIGIN
Query Match 35.6%; Score 27.4; DB 9; Length 454;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 40; Conservative 0; Mismatches 24; Indels 0

QY 12 CCAGCANCATTTAGCATCATCTCTGCATGTCAGGTCATTGGAGGAGN
|||||
395 CCAACCATAGTATCATCAGAGTTTTTCTGCATAGGTAATCTGAACAGAA

Db

QY 72 ANCA 75
|||
335 AAC 332

RESULT 30
AUI72933/6
LOCUS AUI72933 470 bp mRNA linear ES
DEFINITION AUI72933 Rice panicle (between 3cm to 10cm) Oryza sati
cultivar-group) cDNA clone E41160, mRNA sequence.
ACCESSION AUI72933
VERSION AUI72933.1 GI:12622720
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; T
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poac
Ehrhartoideae; Oryzaeae; Oryza.
1 (bases 1 to 470)
Sasaki, T. and Yamamoto, K.
Rice cDNA from panicle (between 3cm to 10cm) (2001)
Unpublished (2001)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.aff
PROJECT = 'RGP'.
Location/Qualifiers
1..470
/organism="Oryza sativa (japonica cultivar-gr
/mol_type="mRNA"

```

```

/cultivar="Nipponbare"
/db xref="taxon:39947"
/clone="E41160"
/tissue type="panicle"
/dev stage="panicle (between 3cm to 10cm)"
/clone_lib="Rice panicle (between 3cm to 10cm)"

35.6%; Score 27.4; DB 9; Length 470;
milarity 62.5%; Pred. No. 1.7e+02;
Conservative 0; Mismatches 24; Indels 0; Gaps 0;

CACACATCTAGCATCATCTGTCATGTCAGGTCATTGGAGGAGNACACAGCC 71
|||||
CAACCATAGTATCATCAGAGTTTCTGCTAGGTAACTGACACAGAACACAGGC 352

INCA 75
|||
LACA 348

487714 482 bp mRNA linear EST 07-FEB-2002
n2n.pk005.g10 Normalized Chicken Breast Muscle, Leg Muscle, and
iphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
one pgm2n.pk005.g10 5' similar to dbj|BAB26278.1 (AK009426)
native [Mus musculus], mRNA sequence.
487714
487714.1 GI:18608644
T.
llus gallus (chicken)
llus gallus
karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
chosauria; Aves; Neognathae; Galliformes; Phasianidae;
asianinae; Gallus.
(bases 1 to 482)
gburn,L.A. and Monsonego-Ornan,E.
ts from Normalized Chicken Breast Muscle, Leg Muscle, and
iphyseal Growth Plate cDNA library, USDA/IFAPS Animal Genome
object
published (2002)
ntact: Larry A. Coghurn
iversity of Delaware
wmsend Hall, Newark, DE 19717, USA
al: 302-831-1335
ax: 302-831-2822
mail: coghurn@udel.edu, www.chickest.udel.edu.
Location/Qualifiers
1. 482
/organism="Gallus gallus"
/mol_type="mRNA"
/strains="Commercial broiler and Ottawa Res. Centre
Strains 90 & 21"
/db xref="taxon:9031"
/clone="pgm2n.pk005.g10"
/sex="Male and Female"
/tissue type="Breast muscle, leg muscle and epiphyseal
growth plate"
/dev stage="Breast, leg: Embryo(d19); post-hatch (1d,1,3,5,7,9
,11 weeks); growth plate(1d,7d,14d post-hatch)"
/lab host="E. coli EMDH10B"
/clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
and Epiphyseal Growth Plate cDNA library (pgm2n)"
/notes="Vector: pCMVSPORT6; Library made from equivalent
pools of total RNA isolated from each tissue (embryonic
muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
plate 33.3% of the final RNA pool). Single pass sequencing
from 5'-end"

35.6%; Score 27.4; DB 12; Length 482;
Similarity 62.5%; Pred. No. 1.7e+02;

```

```

Matches 40; Conservative 0; Mismatches 24; Indels 0;

QY 6 CACAGCCAGCAGCAGCATCTAGCATCATCTCTCCATGGTCAGTCATTGGAG
|||||
DB 282 CTTAGCCCTGCACAGTCACAGCAACATCTGCTTAATGCTGTTGGCATTGGGA
|||||
QY 66 ACAG 69
|||
DB 222 GCTG 219

RESULT 32
CB685142 687 bp mRNA linear EST
OSUNE15F03.r OSUNEF Oryza sativa (japonica cultivar-g:
clone OSUNE15F03 3', mRNA sequence.
CB685142
CB685142.1 GI:29688867
EST.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; T.
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poac.
Ehretaceae; Oryzaceae; Oryza.
1 (bases 1 to 687)
Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou,
Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.
Large-scale identification of ESTs involved in the int-
between rice and Magnaporthe grisea
Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gta g
BACKWARD: gga aac agc tat gac cat g
Plate: 15 row: F column: 03
Seq primer: gga aac agc tat gac cat g.
Location/Qualifiers
1. 687
/organism="Oryza sativa (japonica cultivar-g:
/mol_type="mRNA"
/cultivar="Nipponbare"
/db xref="taxon:39947"
/clone="OSUNE15F03"
/tissue type="Leaf"
/dev stage="3 week"
/lab_host="DH10B"
/clone_lib="OSUNEF"
/notes="Vector: pBluescript II KS +; Site_1:
XhoI; Uninfected Control"

ORIGIN
Query Match 35.6%; Score 27.4; DB 14; Length 687
Best Local Similarity 62.5%; Pred. No. 2e+02;
Matches 40; Conservative 0; Mismatches 24; Indels

QY 12 CCAGCAGCATCTAGCATCATCTCTGTCAGTCAGTCATTGGAGGAGN
|||||
DB 25 CCAACCATAGTATCATCTAGTTTCTGTCATAGGTAACTGACACAGG
|||||
QY 72 ANCA 75
|||
DB 85 AACA 88

RESULT 33
CB474611

```

474611 782 bp DNA linear GSS 27-SEP-2003  
 gr-gss-dog-17000364151175 Dog Library Canis familiaris genomic,  
 nomic survey sequence.  
 474611  
 474611.1 GI:36778061  
 18.  
 unis familiaris (dog)  
 unis familiaris  
 karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 umalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 (bases 1 to 782)  
 rkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,  
 sch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and  
 nter,J.C.  
 he dog genome: survey sequencing and comparative analysis  
 :ence 301 (5641), 1898-1903 (2003)  
 875432  
 512627  
 ntact: Kirkness EF  
 e Institute for Genomic Research  
 partment of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
 ckville, MD 20850, USA  
 i: 301-838-0200  
 x: 301-838-0208  
 ail: ekirknes@tigr.org  
 ass: shotgun.  
 Location/Qualifiers  
 1. 782  
 /organism="Canis familiaris"  
 /mol\_type="genomic DNA"  
 /strain="Standard Poodle"  
 /db\_xref="taxon:9615"  
 /clone\_lib="Dog Library"  
 /note="Site 1: BstXI; Libraries were prepared from  
 peripheral blood"

35.6%; Score 27.4; DB 29; Length 782;  
 imilarity 62.5%; Pred. No. 2.1e+02;  
 ; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
 CCAGCAGCATTCATCATCCTCTGTCATGTCAGGTCATTTGGAGGAGNACAACAGC 70  
 |||||  
 CTCTGGCAGAGTGTGGCTCCCTCTGCGAGGCCAGCTTGCCTAGAGGACCACACAGC 605  
 CAGC 74  
 CTCC 609

190117 984 bp mRNA linear EST 04-SEP-2002  
 ENCOURT\_8076683 NIH\_MGC\_112 Homo sapiens cDNA clone IMAGE:6089012  
 , mRNA sequence.  
 190117  
 190117.1 GI:22704101  
 T.  
 mo sapiens (human)  
 mo sapiens  
 karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 (bases 1 to 984)  
 H-MGC http://mgc.nci.nih.gov/  
 tional Institutes of Health, Mammalian Gene Collection (MGC)  
 ublished (1999)  
 ntact: Robert Strausberg, Ph.D.  
 ail: cgabbs-r@mail.nih.gov  
 sue Procurement: DCTD/DTF  
 DNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 NA Sequencing by: Agencourt Bioscience Corporation  
 lone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLC2328 row: b column: 21  
 High quality sequence stop: 661.

FEATURES  
 Location/Qualifiers  
 1. 984  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6089012"  
 /tissue\_type="melanotic melanoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_112"  
 /note="Organ: skin; Vector: pOTB7; Site 1: Xf  
 EcoRI; cDNA made by oligo-dT priming. Direct  
 into EcoRI/XhoI sites using the following 5',  
 GGACAGAG(G). Library constructed by Ling Hor  
 laboratory of Gerald M. Rubin (University of  
 Berkeley) using ZAP-cDNA synthesis kit (Strat  
 Superscript II RT (Life Technologies)). Note:  
 NIH\_MGC Library."

ORIGIN  
 Query Match 35.6%; Score 27.4; DB 13; Length 984;  
 Best Local Similarity 65.8%; Pred. No. 2.3e+02;  
 Matches 52; Conservative 0; Mismatches 25; Indels ;  
 QY 1 ACTTACNAGCCAGCAGCAGCATTCATCATCCTCTGTCATG-GTCAGGT  
 |||||  
 Db 791 ACTTACTCAGCCAGCATCATTCCTCTCTGGCAGAGTGTCTAGCC  
 QY 59 GAGNACAACAGCCAGCAGT 77  
 |||||  
 Db 851 GAGTACACAGCCATCAGT 869

RESULT 35  
 BM557397  
 LOCUS  
 DEFINITION  
 5', mRNA sequence.  
 BM557397  
 BM557397.1 GI:18799358  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; E  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; ;  
 NIH-MGC http://mgc.nci.nih.gov/  
 1. (bases 1 to 1068)  
 National Institutes of Health, Mammalian Gene Collecti  
 Unpublished (1999)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: ATCC/DCTD/DTF  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (I  
 Clone distribution: MGC clone distribution informati  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM12255 row: n column: 11  
 High quality sequence stop: 643.  
 Location/Qualifiers  
 1. 1068  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5547850"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_72"  
 /note="Organ: skin; Vector: pCMV-SPORT6; Site

FEATURES  
 source  
 Location/Qualifiers  
 1. 1068  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5547850"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_72"  
 /note="Organ: skin; Vector: pCMV-SPORT6; Site

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

35.6%; Score 27.4; DB 12; Length 1068;  
Similarity 65.8%; Pred. No. 2.4e+02;  
; Conservative 0; Mismatches 25; Indels 2; Gaps 1;  
ACTTACNAGCCAGCAGCAGT 77  
|||||  
ACTTACTCAGCCAGCAGCAGT 926  
|||||

3AGNACACAGCAGCAGT 77  
|||||  
3AGTACACAGCAGCAGT 926  
|||||

595860 462 bp mRNA linear EST 25-FEB-2002  
000687477444 A.Gam.ad.cdNA.bloodi Anopheles gambiae cdNA clone  
600449734692 5', mRNA sequence.

595860  
595860.1 GI:18891725

T.  
Anopheles gambiae (African malaria mosquito)

ophelies gambiae  
karyota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
optera; Endopterygota; Diptera; Nematocera; Culicoidea;  
ophelies.

(bases 1 to 462)  
lt.R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L.,  
ariab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.

laria Anopheles gambiae EST project

ublished (2002)

ntact: Holt R.A.

lera Genomics

W. Gude Dr., Rockville, MD 20850, USA

l: 2404533151

lx: 2404534580

mail: HoltR@celera.com

late: NU010049UM row: N column: 22

sq primer: M13 Reverse.

Location/Qualifiers

1..462

/organism="Anopheles gambiae"

/mol\_type="mRNA"

/strain="RSP-ST (Reduced susc. to Permethrin - std.

chromosome)"

/db\_xref="taxon:7165"

/clone="19600449734692"

/dev\_stage="Adult"

/lab\_host="DH10b"

/clone\_lib="A.Gam.ad.cdNA.bloodi"

/note="Vector: pSport1; Site 1: SalI; Site 2: NotI; Whole  
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24  
hours after human blood feeding. cDNA inserts >500 bp  
cloned directionally into pSport 1. Not 1 site is 3'.

Clones available through the Malaria Research and  
Reference Reagent Resource Center (www.malaria.mr4.org)"

35.3%; Score 27.2; DB 12; Length 462;

Similarity 57.9%; Pred. No. 1.9e+02;

4; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

ACTTACNAGCCAGCAGCAGTCTAGCATCCTCTGTCATGCTGTCAGTCAATTCGAGGA 60

|||||  
; AGTTTCGAGCTGATCCCTTTTCGGCAGTAGCTTCTGCGTGGTCAAAATCAATCGGGGA 277

GNACACAGCCAGCAG 76

|||||

Db 276 GCACTTCGGCAGCAG 261

RESULT 37

BM644507/c

LOCUS

DEFINITION 17000687314460 A.Gam.ad.cdNA1 Anopheles gambiae cdNA c

ACCESSION

BM644507

VERSION

BM644507.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..465

/organism="Anopheles gambiae"

/mol\_type="mRNA"

/strain="RSP-ST (Reduced susc. to Permethrin

chromosome)"

/db\_xref="taxon:7165"

/clone="19600449634703"

/dev\_stage="Adult"

/lab\_host="DH10b"

/clone\_lib="A.Gam.ad.cdNA1"

/note="Vector: pSport1; Site 1: SalI; Site 2:

adult mosquitoes (mixed sex) frozen on liquid

cDNA inserts >500 bp cloned directionally in

Not 1 site is 3'. Clones available through

Research and Reference Reagent Resource Cen

(www.malaria.mr4.org)."

Query Match 35.3%; Score 27.2; DB 12; Length 465

Best Local Similarity 57.9%; Pred. No. 1.9e+02;

Matches 44; Conservative 0; Mismatches 32; Indels

QY 1 ACTTACNAGCCAGCAGCAGCAGTCTAGCATCCTCTGTCATGCTGTCAGTCA

Db 306 AGTTTCGAGCTGATCCCTTTTCGGCAGTAGCTTCTGCGTGGTCAATCA

QY 61 GNACACAGCAGCAGCAG 76

Db 246 GCACTTCGGCAGCAG 231

RESULT 38

BM648778/c

LOCUS

DEFINITION 17000687367686 A.Gam.ad.cdNA1 Anopheles gambiae cdNA

19600449621320 5', mRNA sequence.

ACCESSION

BM648778

VERSION

BM648778.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..481

/organism="Anopheles gambiae"

/mol\_type="mRNA"

/strain="RSP-ST (Reduced susc. to Permethrin

chromosome)"

/db\_xref="taxon:7165"

/clone="19600449621320 5', mRNA sequence.

/dev\_stage="Adult"

/lab\_host="DH10b"

/clone\_lib="A.Gam.ad.cdNA1"

/note="Vector: pSport1; Site 1: SalI; Site 2:

adult mosquitoes (mixed sex) frozen on liquid

cDNA inserts >500 bp cloned directionally in

Not 1 site is 3'. Clones available through

Research and Reference Reagent Resource Cen

(www.malaria.mr4.org)."

Db 276 GCACTTCGGCAGCAG 261

RESULT 37

BM644507/c

LOCUS

DEFINITION 17000687314460 A.Gam.ad.cdNA1 Anopheles gambiae cdNA c

ACCESSION

BM644507

VERSION

BM644507.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..465

/organism="Anopheles gambiae"

/mol\_type="mRNA"

/strain="RSP-ST (Reduced susc. to Permethrin

chromosome)"

/db\_xref="taxon:7165"

/clone="19600449634703"

/dev\_stage="Adult"

/lab\_host="DH10b"

/clone\_lib="A.Gam.ad.cdNA1"

/note="Vector: pSport1; Site 1: SalI; Site 2:

adult mosquitoes (mixed sex) frozen on liquid

cDNA inserts >500 bp cloned directionally in

Not 1 site is 3'. Clones available through

Research and Reference Reagent Resource Cen

(www.malaria.mr4.org)."

Query Match 35.3%; Score 27.2; DB 12; Length 465

Best Local Similarity 57.9%; Pred. No. 1.9e+02;

Matches 44; Conservative 0; Mismatches 32; Indels

QY 1 ACTTACNAGCCAGCAGCAGCAGTCTAGCATCCTCTGTCATGCTGTCAGTCA

Db 306 AGTTTCGAGCTGATCCCTTTTCGGCAGTAGCTTCTGCGTGGTCAATCA

QY 61 GNACACAGCAGCAGCAG 76

Db 246 GCACTTCGGCAGCAG 231

RESULT 38

BM648778/c

LOCUS

DEFINITION 17000687367686 A.Gam.ad.cdNA1 Anopheles gambiae cdNA

19600449621320 5', mRNA sequence.

ACCESSION

BM648778

VERSION

BM648778.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..481

/organism="Anopheles gambiae"

/mol\_type="mRNA"

/strain="RSP-ST (Reduced susc. to Permethrin

chromosome)"

/db\_xref="taxon:7165"

/clone="19600449621320 5', mRNA sequence.

/dev\_stage="Adult"

/lab\_host="DH10b"

/clone\_lib="A.Gam.ad.cdNA1"

/note="Vector: pSport1; Site 1: SalI; Site 2:

adult mosquitoes (mixed sex) frozen on liquid

cDNA inserts >500 bp cloned directionally in

Not 1 site is 3'. Clones available through

Research and Reference Reagent Resource Cen

(www.malaria.mr4.org)."

/clone="FK0AAA15AA04"			
/plasmid="pME18S-FL"			
/note="end : 5-PRIME"			
ORIGIN			
Query Match	35.3%	Score 27.2; DB 11; Length 482;	
Best Local Similarity	57.9%	Pred. No. 2e+02;	
Matches 44; Conservative	0; Mismatches 32; Indels 0;		
QY 1 ACTTACNAGCCCGACGANCATTCTAGCATCATCTCTGCATGGTCAGGTCAT			
DB	280 AGTTTCGCAGCTGATCCCTTTCCGCGCAGTAGCTTCTGCGTGGTCAAAATCAAT		
QY 61 GNACACAGCCANCAG 76			
DB	220 GCAACTTCGCAGCAG 205		
RESULT 40			
BM643718/c			
LOCUS	487 bp	mRNA	linear EST
DEFINITION	17000697312981 A.Gam.ad.cdNAL Anopheles gambiae cdNA cl		
ACCESSION	19600449660604 5', mRNA sequence.		
VERSION	BM643718		
KEYWORDS	BM643718.1 GI:18943229		
SOURCE	EST.		
ORGANISM	Anopheles gambiae (African malaria mosquito)		
REFERENCE	Anopheles gambiae		
AUTHORS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pter		
TITLE	Neoptera; Endopterygota; Diptera; Nematocera; Culicoide		
JOURNAL	Anopheles.		
COMMENT	1 (bases 1 to 487)		
	Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.		
	Charlab, R., Collins, F.H., Venter, J.C. and Hoffman, S.L.		
	Celera Anopheles gambiae EST project		
	Unpublished (2002)		
	Contact: Holt R.A.		
	Celera Genomics		
	45 W. Gude Dr., Rockville, MD 20850, USA		
	Tel: 2404533151		
	Fax: 2404534580		
	Email: Holtra@celera.com		
	Plate: NU01004HNP row: P column: 10		
	Seq primer: M13 Reverse.		
FEATURES			
source	Location/Qualifiers		
	1..487		
	/organism="Anopheles gambiae"		
	/mol_type="mRNA"		
	/strain="RSP-ST (Reduced susc. to Permethrin -		
	chromosome)"		
	/db_xref="taxon:7165"		
	/clone="19600449660604"		
	/dev_stage="Adult"		
	/lab_host="DH10b"		
	/clone_lib="A.Gam.ad.cdNAL"		
	note=Vector: pSport1; Site 1: SalI; Site 2:		
	adult mosquitoes (mixed sex) frozen on liquid		
	cDNA inserts >500 bp cloned directionally int		
	Not 1 site is 3'. Clones available through th		
	Research and Reference Reagent Resource Cente		
	(www.malaria.mr4.org)."		
ORIGIN			
Query Match	35.3%	Score 27.2; DB 12; Length 487;	
Best Local Similarity	57.9%	Pred. No. 2e+02;	
Matches 44; Conservative	0; Mismatches 32; Indels 0;		
QY 1 ACTTACNAGCCCGACGANCATTCTAGCATCATCTCTGCATGGTCAGGTCAT			
DB	327 AGTTTCGCAGCTGATCCCTTTCCGCGCAGTAGCTTCTGCGTGGTCAAAATCAAT		
QY 61 GNACACAGCCANCAG 76			

.4 08:44:47 2004

us-10-090-326-25.rst

30RACCTCGGACGAG 252

ad: April 13, 2004, 20:14:23  
7.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

ncleic search, using sw model

April 13, 2004, 19:12:18 ; Search time 174.5 Seconds  
(without alignments)  
1655.373 Million cell updates/sec

US-10-090-326-25

1 acttaacagccagcanca.....ggagnacaacagcancagt 77

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

2475585 seqs, 1875730760 residues

: hits satisfying chosen parameters: 4951170

length: 0

length: 2000000000

3: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Published Applications NA:

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

is the number of results predicted by chance to have a  
water than or equal to the score of the result being printed,  
derived by analysis of the total score distribution.

#### SUMMARIES

Query Match	Length	ID	Description
39.0	1059	10	US-09-940-925A-58
39.0	1059	10	US-09-940-925A-58
39.0	1059	10	US-09-941-193A-58
39.0	1059	10	US-09-941-193A-59
39.0	1587	10	US-09-940-925A-60
39.0	1587	10	US-09-940-925A-61
39.0	1587	10	US-09-941-193A-60
39.0	1587	10	US-09-941-193A-61
39.0	1906	13	US-10-011-436-1
39.0	1964	15	US-10-117-937-5
39.0	2041	9	US-09-834-975-959
39.0	6408	9	US-09-996-128A-1
38.2	918	15	US-10-027-632-30884
38.2	918	15	US-10-027-632-30885
36.9	100	14	US-10-057-810-41

16	28.4	36.9	100	14	US-10-057-828-41	Seque
17	28.4	36.9	3075	15	US-10-369-493-24411	Seque
18	28.4	36.9	3078	15	US-10-410-012-8	Seque
19	28.4	36.9	3150	14	US-10-162-688-1	Seque
20	28.4	36.9	3157	9	US-09-965-703-74	Sequen
21	28.4	36.9	3284	14	US-10-157-305A-186	Seque
22	28.4	36.9	3284	14	US-10-157-391-186	Seque
23	28.4	36.9	3284	14	US-10-157-096-186	Seque
24	28.4	36.9	3284	14	US-10-157-302-186	Seque
25	28.4	36.9	3284	14	US-10-157-215A-186	Seque
26	28.4	36.9	3284	14	US-10-157-299-186	Seque
27	28.4	36.9	3284	14	US-10-154-951B-186	Seque
28	28.4	36.9	3284	14	US-10-156-831-186	Seque
29	28.4	36.9	3284	14	US-10-157-147-186	Seque
30	28.4	36.9	3284	14	US-10-157-166-186	Seque
31	28.4	36.9	3284	14	US-10-156-902-186	Seque
32	28.4	36.9	3284	14	US-10-157-318-186	Seque
33	28.4	36.9	3284	15	US-10-156-811-186	Seque
34	28.4	36.9	3284	15	US-10-157-320A-186	Seque
35	28.4	36.9	3284	15	US-10-157-418A-186	Seque
36	28.4	36.9	3284	15	US-10-157-317-186	Seque
37	28.4	36.9	3284	15	US-10-157-339-186	Seque
38	28.4	36.9	3396	14	US-10-338-411-32	Seque
39	28.4	36.9	3396	15	US-10-389-640-32	Sequen
40	28.4	36.9	4279	9	US-09-956-988A-1	Sequen
41	28.4	36.9	5646	12	US-10-270-176-39	Sequen
42	28.4	36.9	5646	12	US-10-270-176-41	Sequen
43	28.4	36.9	5706	12	US-10-270-176-40	Sequen
44	28.4	36.9	5739	12	US-10-270-176-9	Sequen
45	28.4	36.9	5750	12	US-10-270-176-32	Sequen

#### ALIGNMENTS

#### RESULT 1

US-09-940-925A-58  
; Sequence 58, Application US/09940925A  
; Publication No. US20030054338A1  
; GENERAL INFORMATION:  
; APPLICANT: BROW, MARY ANN D.  
; OLIVE, DAVID M.  
; LYAMICHEV, VICTOR I.  
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
; NUMBER OF SEQUENCES: 165  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MEDLEN & CARROLL  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/940,925A  
; FILING DATE: 10-Jun-2002  
; CLASSIFICATION: <UNKNOWN>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CARROLL, PETER G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: FORS-01756  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 58:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1059 base pairs  
; TYPE: nucleic acid

```

Db      535 AGTACACAGCCATCAGT 552

RESULT 3
US-09-941-193A-58
; Sequence 58, Application US/09941193A
; Publication No. US20030108873A1
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
;             LYAMICHEV, VICTOR I.
;             OLIVE, DAVID M.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/941,193A
; FILING DATE: 28-Aug-2001
; CLASSIFICATION: <UNKNOWN>
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1059 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-09-941-193A-58

Query Match      39.0%; Score 30; DB 10; Length 1059;
Best Local Similarity 67.9%; Pred No. 0.17; Mismatches 1
Matches 53; Conservative 0;

QY      1 ACTTACNAGCCCGACGANCATCTAGCATCATCTCTGCATG-GTCAGGTCA
Db      475 ACTTACTCAGCCCGACGATCATCTCTCTCTTGGCAGATTCTCTGTAGCCG
QY      60 AGNACAACAGCCNACGT 77
Db      535 AGTACACAGCCATCAGT 552

RESULT 4
US-09-941-193A-59
; Sequence 59, Application US/09941193A
; Publication No. US20030108873A1
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
;             LYAMICHEV, VICTOR I.
;             OLIVE, DAVID M.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL

```



TREET: 220 MONTGOMERY STREET, SUITE 2200  
ITY: SAN FRANCISCO  
TATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
IP: 94104  
ER READABLE FORM:  
EDM TYPE: Floppy disk  
MPUTER: IBM PC compatible  
PERATING SYSTEM: PC-DOS/MS-DOS  
FTWARE: Patentin Release #1.0, Version #1.30  
T APPLICATION DATA:  
PLICATION NUMBER: US/09/941.193A  
ILING DATE: 28-Aug-2001  
CLASSIFICATION: <Unknown>  
EY/AGENT INFORMATION:  
AME: CARROLL, PETER G.  
EGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: FORS-01756  
MMUNICATION INFORMATION:  
ELEPHONE: (415) 705-8410  
ELEFAX: (415) 397-8338  
FOR SEQ ID NO: 59:  
CE CHARACTERISTICS:  
ENGTH: 1059 base pairs  
YPE: nucleic acid  
TRANDEDNESS: double  
POLOGY: linear  
LE TYPE: DNA (genomic)  
CE DESCRIPTION: SEQ ID NO: 59:  
59  
39.0%; Score 30; DB 10; Length 1059;  
nilarity 67.9%; Pred. No. 0.17; 24; Indels 1; Gaps 1;  
Conservative 0; Mismatches 0; Mismatches 24; Indels 1;  
TTACNAGCCAGCCAGCATTCTAGCATCTCTGCGATG-GTCAGGTCATTGGAGG 59  
TTTACTCAGCCAGCATCATCTTCTCTCTTGGCAGATTGCTGTAGCCGATTGGAGG 534  
TACACAGCCANCACT 77  
TACACAGCCATCACT 552  
60  
Application US/09940925A  
2. US20030054338A1  
ORMATION:  
ANT: BROW, MARY ANN D.  
LYAMICHEV, VICTOR I.  
OLIVE, DAVID M.  
OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
PATHOGENS  
OF SEQUENCES: 165  
PONDENCE ADDRESS:  
DRESSEE: MEDLEN & CARROLL  
TREET: 220 MONTGOMERY STREET, SUITE 2200  
ITY: SAN FRANCISCO  
TATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
IP: 94104  
ER READABLE FORM:  
EDM TYPE: Floppy disk  
MPUTER: IBM PC compatible  
PERATING SYSTEM: PC-DOS/MS-DOS  
FTWARE: Patentin Release #1.0, Version #1.30  
T APPLICATION DATA:  
PLICATION NUMBER: US/09/940.925A  
ILING DATE: 10-Jun-2002  
CLASSIFICATION: <Unknown>  
EY/AGENT INFORMATION:  
AME: CARROLL, PETER G.

REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: FORS-01756  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1587 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 60:  
US-09-940-925A-60  
Query Match 39.0%; Score 30; DB 10; Length 1587;  
Best Local Similarity 67.9%; Pred. No. 0.18;  
Matches 53; Conservative 0; Mismatches 24; Indels 1;  
QY 1 ACTTACNAGCCAGCCAGCATTCTAGCATCTCTGCGATG-GTCAGGTCAT  
Db 782 ACTTACTCAGCCAGCATCATCTTCTCTCTTGGCAGATTGCTGTAGCCGA  
QY 60 AGNACAACAGCCANCACT 77  
Db 842 AGTACACAGCCATCACT 859  
RESULT 6  
US-09-940-925A-61  
Sequence 61, Application US/09940925A  
Publication No. US20030054338A1  
GENERAL INFORMATION:  
APPLICANT: BROW, MARY ANN D.  
LYAMICHEV, VICTOR I.  
OLIVE, DAVID M.  
TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
PATHOGENS  
NUMBER OF SEQUENCES: 165  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MEDLEN & CARROLL  
STREET: 220 MONTGOMERY STREET, SUITE 2200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/940.925A  
FILING DATE: 10-Jun-2002  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: CARROLL, PETER G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: FORS-01756  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1587 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 61:  
US-09-940-925A-61  
Query Match 39.0%; Score 30; DB 10; Length 1587;

milarity 67.9%; Pred. No. 0.18; DB 10; Length 1587;  
 Conservative 0; Mismatches 24; Indels 1; Gaps 1;  
 CTATACAGCCAGCAGCATCTTCTGCGATG-GTCAGGTCATTGGAGG 59  
 |||||  
 CTATCTCAGCCAGCATCTTCTCTCTTGGCAGATTGCTGTAGCCGATTGGAGG 841  
 |||||  
 GNACACAGCCAGCAT 77  
 |||||  
 GTACACAGCCATCAGT 859  
 |||||  
 Application US/09941193A  
 O. US20030108873A1  
 INVENTOR:  
 BROW, MARY ANN D.  
 LYAMICHEV, VICTOR I.  
 OLIVE, DAVID M.  
 OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
 PATHOGENS  
 OF SEQUENCES: 165  
 PONDENCE ADDRESS:  
 DRESSEE: MEDLEN & CARROLL  
 TREET: 220 MONTGOMERY STREET, SUITE 2200  
 ITY: SAN FRANCISCO  
 COUNTRY: UNITED STATES OF AMERICA  
 IP: 94104  
 ER READABLE FORM:  
 EDIUM TYPE: Floppy disk  
 MPUTER: IBM PC compatible  
 PERATING SYSTEM: PC-DOS/MS-DOS  
 FTWARE: Patent in Release #1.0, Version #1.30  
 T APPLICATION DATA:  
 PPLICATION NUMBER: US/09/941,193A  
 ILING DATE: 28-Aug-2001  
 ASSIFICATION: <Unknown>  
 EX/AGENT INFORMATION:  
 AME: CARROLL, PETER G.  
 EGISTRATION NUMBER: 32,837  
 REFERENCE/DOCKET NUMBER: FORS-01756  
 MMUNICATION INFORMATION:  
 ELEPHONE: (415) 705-8410  
 ELEFAX: (415) 397-8338  
 FOR SEQ ID NO: 60:  
 CE CHARACTERISTICS:  
 ENGT: 1587 base pairs  
 YPE: nucleic acid  
 TRANDEDNESS: double  
 OPOLOGY: linear  
 LE TYPE: DNA (genomic)  
 CE DESCRIPTION: SEQ ID NO: 60:  
 60  
 milarity 39.0%; Score 30; DB 10; Length 1587;  
 Conservative 0; Mismatches 24; Indels 1; Gaps 1;  
 CTATACAGCCAGCAGCATCTTCTGCGATG-GTCAGGTCATTGGAGG 59  
 |||||  
 CTATCTCAGCCAGCATCTTCTCTCTTGGCAGATTGCTGTAGCCGATTGGAGG 841  
 |||||  
 GNACACAGCCAGCAT 77  
 |||||  
 GTACACAGCCATCAGT 859  
 |||||  
 Application US/09941193A  
 O. US20030108873A1

GENERAL INFORMATION:  
 APPLICANT: BROW, MARY ANN D.  
 LYAMICHEV, VICTOR I.  
 OLIVE, DAVID M.  
 TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
 PATHOGENS  
 NUMBER OF SEQUENCES: 165  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MEDLEN & CARROLL  
 STREET: 220 MONTGOMERY STREET, SUITE 2200  
 CITY: SAN FRANCISCO  
 STATE: CALIFORNIA  
 COUNTRY: UNITED STATES OF AMERICA  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/941,193A  
 FILING DATE: 28-Aug-2001  
 CLASSIFICATION: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: CARROLL, PETER G.  
 REGISTRATION NUMBER: 32,837  
 REFERENCE/DOCKET NUMBER: FORS-01756  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 705-8410  
 TELEFAX: (415) 397-8338  
 INFORMATION FOR SEQ ID NO: 61:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1587 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 SEQUENCE DESCRIPTION: SEQ ID NO: 61:  
 US-09-941-193A-61  
 Query Match 39.0%; Score 30; DB 10; Length 1587;  
 Best Local Similarity 67.9%; Pred. No. 0.18;  
 Matches 53; Conservative 0; Mismatches 24; Indels 1;  
 QY 1 ACTTACAGCCAGCAGCATCTTCTGCGATG-GTCAGGTCAT  
 |||||  
 Db 782 ACTTACTCAGCCAGCATCTTCTCTCTTGGCAGATTGCTGTAGCCGA  
 |||||  
 QY 60 AGNACACAGCCAGCAT 77  
 |||||  
 Db 842 AGTACACAGCCATCAGT 859  
 |||||  
 RESULT 9  
 US-10-011-436-1  
 ; Sequence 1, Application US/10011436  
 ; Publication No. US20020128200A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Boon-Falleur, Thierry; Brichard, Vincent; Van P  
 ; Aline;  
 ; Deplaen, Etienne; Coullie, Pierre; Renauld, Jean  
 ; Lerche, Bernard  
 ; TITLE OF INVENTION: METHOD FOR IDENTIFYING INDIVIDUALS SUF  
 ; A  
 ; CELLULAR ABNORMALITY SOME OF WHOSE ABN  
 ; LEUKOCYTE  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fulbright & Jaworski L.L.P.  
 ; STREET: 801 Pennsylvania Avenue N.W.  
 ; CITY: Washington  
 ; STATE: District of Columbia  
 ; COUNTRY: USA

IP: 20004  
 ER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.25inch, 1.44MB storage  
 COMPUTER: IBM PS/2  
 OPERATING SYSTEM: PC-DOS  
 SOFTWARE: Wordperfect  
 APPLICATION DATA:  
 APPLICATION NUMBER: US/10/011,436  
 FILING DATE: 11-Dec-2001  
 CLASSIFICATION: 435  
 APPLICATION DATA:  
 APPLICATION NUMBER: 08/511,011  
 FILING DATE: 5-AUGUST-1995  
 APPLICATION NUMBER: 08/054,714  
 FILING DATE: 28-APRIL-1993  
 APPLICATION NUMBER: 07/994,928  
 FILING DATE: 22-DECEMBER-1992  
 KEY/AGENT INFORMATION:  
 NAME: Mary Anne Schofield  
 REGISTRATION NUMBER: 36,669  
 REFERENCE/DOCKET NUMBER: LUD 5299.9 DIV CON  
 COMMUNICATION INFORMATION:  
 TELEPHONE: (202) 662-0200  
 TELEFAX: (202) 662-4643  
 FOR SEQ ID NO: 1:  
 CE CHARACTERISTICS:  
 LENGTH: 1906 base pairs  
 TYPE: nucleic acid  
 TRANSDENESS: single  
 TOPOLOGY: linear  
 ICE DESCRIPTION: SEQ ID NO: 1:

39.0%; Score 30; DB 13; Length 1906;  
 milarity 67.9%; Pred. No. 0.19;  
 Conservative 0; Mismatches 24; Indels 1; Gaps 1;  
 CTTACNAGCCGAGCAGCATCTTCTGATCATCTCTGATG-CTCAGGTCATTGGAGG 59  
 CTTACTCAGCCGAGCAGCATCTTCTCTTGGCAGATTGCTGTAGCGGATTGGAGG 865

GNACACAGCCANCACT 77  
 |||||  
 GTACACAGCCATCACT 883

pplication US/10117937  
 O US2003020239A1  
 NATION:  
 TL IMMUNO THERAPIES CORP.  
 SIMARD, John, J.L.  
 DIAMOND, David, C.  
 Liu, Liping  
 XIE, Zhidong  
 ENTION: EPTOPE SEQUENCES  
 CE: CTLIMM.027A  
 ICATION NUMBER: US/10/117,937  
 NG DATE: 2002-04-04  
 ATION NUMBER: US 60/282,211  
 DATE: 2001-04-06  
 ATION NUMBER: US 60/337,017  
 DATE: 2001-11-07  
 ATION NUMBER: US 60/363,210  
 DATE: 2002-03-07  
 Q ID NOS: 602  
 st-SEQ for Windows Version 4.0

4

omo sapiens

Query Match 39.0%; Score 30; DB 15; Length 1964;  
 Best Local Similarity 67.9%; Pred. No. 0.19;  
 Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNAGCCGAGCAGCATCTTCTGATCATCTCTGATG-CTCAGGTCAT  
 |||||  
 Db 864 ACTTACTCAGCCGAGCAGCATCTTCTCTCTTGGCAGATTGCTGTAGCCG  
 |||||  
 Qy 60 AGNACACAGCCANCACT 77  
 |||||  
 Db 924 AGTACACAGCCATCACT 941  
 |||||

RESULT 11  
 US-09-834-975-959  
 ; Sequence 959, Application US/09834975  
 ; Patent No. US20020110815A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lillie, James  
 ; APPLICANT: Brown, Jeffrey  
 ; APPLICANT: Bolt, Andrew  
 ; APPLICANT: Van Huffel, Christophe  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS  
 ; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENT  
 ; FILE REFERENCE: MRI-016B  
 ; CURRENT APPLICATION NUMBER: US/09/834,975  
 ; CURRENT FILING DATE: 2001-04-13  
 ; PRIOR APPLICATION NUMBER: 60/197,538  
 ; PRIOR FILING DATE: 2000-04-14  
 ; NUMBER OF SEQ ID NOS: 1046  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 959  
 ; LENGTH: 2041  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)...(2041)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-834-975-959

Query Match 39.0%; Score 30; DB 9; Length 2041;  
 Best Local Similarity 67.9%; Pred. No. 0.2;  
 Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNAGCCGAGCAGCATCTTCTGATCATCTCTGATG-CTCAGGTCAT  
 |||||  
 Db 907 ACTTACTCAGCCGAGCAGCATCTTCTCTCTTGGCAGATTGCTGTAGCCGA  
 |||||  
 Qy 60 AGNACACAGCCANCACT 77  
 |||||  
 Db 967 AGTACACAGCCATCACT 984  
 |||||

RESULT 12  
 US-09-996-128A-1  
 ; Sequence 1, Application US/09996128A  
 ; Patent No. US20020150589A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Houghton, Alan  
 ; APPLICANT: Bergman, Phillip  
 ; APPLICANT: Wolchok, Jedd  
 ; TITLE OF INVENTION: Compositions for treatment of Melanoma and M  
 ; FILE REFERENCE: MSK P-026-3  
 ; CURRENT APPLICATION NUMBER: US/09/996,128A  
 ; CURRENT FILING DATE: 2001-11-27  
 ; PRIOR APPLICATION NUMBER: US 09/627,694  
 ; PRIOR FILING DATE: 2000-07-28  
 ; PRIOR APPLICATION NUMBER: US 09/308,697  
 ; PRIOR FILING DATE: 1999-05-21  
 ; PRIOR APPLICATION NUMBER: PCT/US97/22669

HCACAGCCAGGCTAGATTATCTCGTGTGTC

Best Local Similarity 96.

ACACAGCCAGGCTAGATTATCTCGYTGTCATAATCTTTGCAAGGCAGTTTCAGGAGAAC 248

Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
TAGCATCATCTCTGCATGTCAGGTCAT 52  
|||||  
GAGCATCATCTCTGCATGTCAGGTCAT 76  
|||||  
1 Application US/10057828  
O. US20030148287A1  
NATION:  
i, Xiangqiang  
Jiang, Xin  
ENTION: LIBRARIES AND KITS FOR DETECTING TRANSCRIPTION FACTOR ACTIVITY  
CE: 26757-707  
ICATION NUMBER: US/10/057,828  
NG DATE: 2002-01-24  
Q ID NOS: 60  
tentIn version 3.1  
rtificial Sequence  
MATION: Pp11 Reporter Sequence  
I  
milarity 36.9%; Score 28.4; DB 14; Length 100;  
milarity 96.7%; Pred. No. 0.37;  
Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
TAGCATCATCTCTGCATGTCAGGTCAT 52  
|||||  
GAGCATCATCTCTGCATGTCAGGTCAT 76  
|||||  
4411  
1, Application US/10369493  
O. US20030233675A1  
MATION:  
ao, Yongwei  
Hinkle, Gregory J.  
Slater, Steven C.  
Goldman, Barry S.  
Chen, Xianfeng  
ENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
ENTION: PLANTS WITH IMPROVED PROPERTIES  
CE: 38-10(52052)B  
ICATION NUMBER: US/10/369,493  
NG DATE: 2003-02-28  
ATION NUMBER: US 60/360,039  
DATE: 2002-02-21  
Q ID NOS: 47374  
11  
5  
scherichia coli  
4411  
milarity 36.9%; Score 28.4; DB 15; Length 3075;  
milarity 96.7%; Pred. No. 0.93;  
Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
TAGCATCATCTCTGCATGTCAGGTCAT 52  
|||||  
GAGCATCATCTCTGCATGTCAGGTCAT 1103  
|||||  
pplication US/10410012

Publication No. US20030228276A1  
; GENERAL INFORMATION:  
; APPLICANT: Pitossi, Fernando J  
; APPLICANT: Eisel, Ulrich L M  
; TITLE OF INVENTION: Neuroprotective and neurodegenerative effect  
; TITLE OF INVENTION: in the substantia nigra and a new model fo  
; FILE REFERENCE: 1418P US/101141-17  
; CURRENT APPLICATION NUMBER: US/10/410,012  
; CURRENT FILING DATE: 2003-04-09  
; PRIOR APPLICATION NUMBER: US 60/370,974  
; PRIOR FILING DATE: 2002-04-09  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-10-410-012-8  
Query Match 36.9%; Score 28.4; DB 15; Length 3078;  
Best Local Similarity 96.7%; Pred. No. 0.93;  
Matches 29; Conservative 0; Mismatches 1; Indels 0;  
QY 23 CTAGCATCATCTCTGCATGTCAGGTCAT 52  
Db 1071 CGAGCATCATCTCTGCATGTCAGGTCAT 1100  
|||||  
RESULT 19  
US-10-162-688-1  
; Sequence 1, Application US/10162688  
; Publication No. US20030100528A1  
; GENERAL INFORMATION:  
; APPLICANT: INSTITUT PASTEUR  
; APPLICANT: CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE  
; TITLE OF INVENTION: ISOLATED POLYNUCLEOTIDES HAVING A REDUCED OF  
; TITLE OF INVENTION: INCREASED CONTENT OF EPIGENETIC CONTROL MOT  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: B4752 IP  
; CURRENT APPLICATION NUMBER: US/10/162,688  
; CURRENT FILING DATE: 2002-06-06  
; PRIOR APPLICATION NUMBER: CANADA NO 2,291,367  
; PRIOR FILING DATE: 1999-12-06  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 3150  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(3150)  
; OTHER INFORMATION: Nucleic acid sequence of LacZ  
US-10-162-688-1  
Query Match 36.9%; Score 28.4; DB 14; Length 3150;  
Best Local Similarity 96.7%; Pred. No. 0.93;  
Matches 29; Conservative 0; Mismatches 1; Indels 0;  
QY 23 CTAGCATCATCTCTGCATGTCAGGTCAT 52  
Db 1149 CGAGCATCATCTCTGCATGTCAGGTCAT 1178  
|||||  
RESULT 20  
US-09-965-703-74  
; Sequence 74, Application US/09965703  
; Patent No. US20020119521A1  
; GENERAL INFORMATION:  
; APPLICANT: Rohm and Haas Company  
; APPLICANT: Palli, Subba Reddy  
; APPLICANT: Kapitskaya, Marianna Zinovjevna  
; APPLICANT: Cress, Dean Ervin

```

ENTION: No. US20020119521A1e1 Ecdysone Receptor-Based Inducible Gene Exp
CE: A01020B
ICATION NUMBER: US/09/965,703
NG DATE: 2001-09-26
ATION NUMBER: 60/191,355
DATE: 2000-03-22
ATION NUMBER: 60/269,799
DATE: 2001-02-20
ATION NUMBER: PCT/US01/09050
DATE: 2001-03-21
Q ID NOS: 75
tentin version 3.1
7
4
Escherichia coli
36.9%; Score 28.4; DB 9; Length 3157;
milarity 96.7%; Pred. No. 0.93;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CTAGCATCATCCTCTGTCATGGTCAGGTCAAT 52
GAGCATCATCCTCTGTCATGGTCAGGTCAAT 1199
186
Application US/10157305A
No. US20030166099A1
INATION:
abbadini, Roger A.
Surber, Mark W.
Neil Berkley
Anca M. Segall
Robert Klepper
VENTION: MINICELL COMPRISING MEMBRANE PROTEINS
ICE: MPEX.008DV1
NG DATE: 2002-05-28
ATION NUMBER: 60/293,566
DATE: 2001-05-24
ATION NUMBER: 60/359,843
DATE: 2002-02-25
ATION NUMBER: 10/154,951
DATE: 2002-05-24
Q ID NOS: 258
astSEQ for Windows Version 4.0
5
34
Artificial Sequence
INATION: Gene encoding a fusion protein
-186
36.9%; Score 28.4; DB 14; Length 3284;
milarity 96.7%; Pred. No. 0.94;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CTAGCATCATCCTCTGTCATGGTCAGGTCAAT 52
CGAGCATCATCCTCTGTCATGGTCAGGTCAAT 1283
186
Application US/10157391
No. US20030166279A1
INATION:
abbadini, Roger A.
Neil Berkley

```

```

; TITLE OF INVENTION: MINICELL-BASED TRANSFECTION
; FILE REFERENCE: MPEX.008DV14
; CURRENT APPLICATION NUMBER: US/10/157,391
; CURRENT FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/359,843
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 10/154,951
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 186
; LENGTH: 3284
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
US-10-157-391-186
Query Match 36.9%; Score 28.4; DB 14; Length 3284;
Best Local Similarity 96.7%; Pred. No. 0.94;
Matches 29; Conservative 0; Mismatches 1; Indels 0;
Qy 23 CTAGCATCATCCTCTGTCATGGTCAGGTCAAT 52
Db 1254 CGAGCATCATCCTCTGTCATGGTCAGGTCAAT 1283
RESULT 23
US-10-157-096-186
; Sequence 186, Application US/10157096
; Publication No. US20030190601A1
; GENERAL INFORMATION:
; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Neil Berkley
; APPLICANT: Surber, Mark W.
; TITLE OF INVENTION: TARGET DISPLAY ON MINICELLS
; FILE REFERENCE: MPEX.008DV12
; CURRENT APPLICATION NUMBER: US/10/157,096
; CURRENT FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/359,843
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 10/154,51
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 186
; LENGTH: 3284
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
US-10-157-096-186
Query Match 36.9%; Score 28.4; DB 14; Length 3284
Best Local Similarity 96.7%; Pred. No. 0.94;
Matches 29; Conservative 0; Mismatches 1; Indels 0.
Qy 23 CTAGCATCATCCTCTGTCATGGTCAGGTCAAT 52
Db 1254 CGAGCATCATCCTCTGTCATGGTCAGGTCAAT 1283
RESULT 24
US-10-157-302-186
; Sequence 186, Application US/10157302
; Publication No. US20030190689A1
; GENERAL INFORMATION:
; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Surber, Mark W.

```

## VENTION: MINICELL-BASED RATIONAL DRUG DESIGN

ACE: MPX.008DV17  
 ICACTION NUMBER: US/10/157,302  
 ING DATE: 2002-10-01  
 ATION NUMBER: 60/293,566  
 ; DATE: 2001-05-24  
 ATION NUMBER: 60/359,843  
 ; DATE: 2002-02-25  
 ATION NUMBER: 10/154,951  
 ; DATE: 2002-05-24  
 Q ID NOS: 257  
 stSEQ for Windows Version 4.0  
 ;  
 14  
 Artificial Sequence  
 MATION: Fusion protein  
 .86  
 milarity 36.9%; Score 28.4; DB 14; Length 3284;  
 Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 TAGCATCATCCTCTGCATGGTCAGGTGCAT 52  
 |||||  
 GAGCATCATCCTCTGCATGGTCAGGTGCAT 1283

186  
 Application US/10157215A  
 o. US20030190749A1  
 MATION:  
 Surber, Mark W.  
 Sabbadini, Roger A.  
 Segall, Anca M.  
 Berkley, Neil  
 VENTION: MINICELL-PRODUCING PARENT CELLS  
 ACE: MPX.008DV23  
 ICACTION NUMBER: US/10/157,215A  
 NG DATE: 2002-05-28  
 ATION NUMBER: 60/293,566  
 ; DATE: 2001-05-24  
 ATION NUMBER: 60/359,843  
 ; DATE: 2002-02-25  
 ATION NUMBER: 10/154,951  
 ; DATE: 2002-05-24  
 Q ID NOS: 258  
 stSEQ for Windows Version 4.0  
 ;  
 4  
 Artificial Sequence  
 MATION: Gene encoding a fusion protein  
 186  
 milarity 36.9%; Score 28.4; DB 14; Length 3284;  
 Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 TAGCATCATCCTCTGCATGGTCAGGTGCAT 52  
 |||||  
 GAGCATCATCCTCTGCATGGTCAGGTGCAT 1283

86  
 Application US/10157299  
 o. US20030194714A1  
 MATION:  
 Sabbadini, Roger A.

; APPLICANT: Neil Berkley  
 ; APPLICANT: Surber, Mark W.  
 ; TITLE OF INVENTION: MINICELL-BASED TRANSFORMATION  
 ; FILE REFERENCE: MPX.008DV15  
 ; CURRENT APPLICATION NUMBER: US/10/157,299  
 ; CURRENT FILING DATE: 2002-10-01  
 ; PRIOR APPLICATION NUMBER: 60/293,566  
 ; PRIOR FILING DATE: 2001-05-24  
 ; PRIOR APPLICATION NUMBER: 60/359,843  
 ; PRIOR FILING DATE: 2002-02-25  
 ; PRIOR APPLICATION NUMBER: 10/154,951  
 ; PRIOR FILING DATE: 2002-05-24  
 ; NUMBER OF SEQ ID NOS: 257  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 186  
 ; LENGTH: 3284  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Fusion protein  
 US-10-157-299-186

Query Match 36.9%; Score 28.4; DB 14; Length 3284  
 Best Local Similarity 96.7%; Pred. No. 0.94;  
 Matches 29; Conservative 0; Mismatches 1; Indels 0  
 Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTGCAT 52  
 |||||  
 Db 1254 CGAGCATCATCCTCTGCATGGTCAGGTGCAT 1283

RESULT 27  
 US-10-154-951B-186  
 ; Sequence 186, Application US/10154951B  
 ; Publication No. US20030194798A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Surber, Mark W.  
 ; APPLICANT: Sabbadini, Roger A.  
 ; TITLE OF INVENTION: MINICELL COMPOSITIONS AND METHODS  
 ; FILE REFERENCE: MPX.008A  
 ; CURRENT APPLICATION NUMBER: US/10/154,951B  
 ; CURRENT FILING DATE: 2002-05-24  
 ; PRIOR APPLICATION NUMBER: 60/293,566  
 ; PRIOR FILING DATE: 2001-05-24  
 ; PRIOR APPLICATION NUMBER: 60/359,843  
 ; PRIOR FILING DATE: 2002-02-25  
 ; NUMBER OF SEQ ID NOS: 258  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 186  
 ; LENGTH: 3284  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Gene encoding a fusion protein  
 US-10-154-951B-186

Query Match 36.9%; Score 28.4; DB 14; Length 3284;  
 Best Local Similarity 96.7%; Pred. No. 0.94;  
 Matches 29; Conservative 0; Mismatches 1; Indels 0;  
 Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTGCAT 52  
 |||||  
 Db 1254 CGAGCATCATCCTCTGCATGGTCAGGTGCAT 1283

RESULT 28  
 US-10-156-831-186  
 ; Sequence 186, Application US/10156831  
 ; Publication No. US20030198995A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sabbadini, Roger A.  
 ; APPLICANT: Neil Berkley  
 ; APPLICANT: Surber, Mark W.

## ENTION: FORWARD SCREENING USING MINICELLS

ICE: MPEX.008DV20  
 ICAION NUMBER: US/10/156,831  
 NG DATE: 2002-10-02  
 ATION NUMBER: 60/293,566  
 ; DATE: 2001-05-24  
 ATION NUMBER: 60/359,843  
 ; DATE: 2002-02-25  
 ATION NUMBER: 10/154,951  
 ; DATE: 2002-05-24  
 ; ID NOS: 257  
 ;stSEQ for Windows Version 4.0

14

## Artificial Sequence

IMATION: Fusion protein

.86

36.9%; Score 28.4; DB 14; Length 3284;  
 milarity 96.7%; Pred. No. 0.94;  
 Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CTAGCATCATCCTCTGCATGGTCAGGTGCAT 52  
 |||||  
 CGAGCATCATCCTCTGCATGGTCAGGTGCAT 1283

.86

Application US/10157147  
 No. US20030198996A1

IMATION:

Surber, Mark W.

Berkley, Neil

Gerhart, William

Sabbadini, Roger A.

ENTION: MINICELL LIBRARIES

ICE: MPEX.008DV19

ICATION NUMBER: US/10/157,147  
 NG DATE: 2002-05-28

ATION NUMBER: 60/293,566  
 ; DATE: 2001-05-24

ATION NUMBER: 60/359,843  
 ; DATE: 2002-02-25

ATION NUMBER: 10/154,951  
 ; DATE: 2002-05-24

; ID NOS: 258

stSEQ for Windows Version 4.0

14

## Artificial Sequence

IMATION: Gene encoding a fusion protein

.86

36.9%; Score 28.4; DB 14; Length 3284;  
 milarity 96.7%; Pred. No. 0.94;  
 Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CTAGCATCATCCTCTGCATGGTCAGGTGCAT 52  
 |||||  
 CGAGCATCATCCTCTGCATGGTCAGGTGCAT 1283

.86

Application US/10157166  
 No. US20030199005A1

IMATION:

Sabbadini, Roger A.

; APPLICANT: Robert Klepper  
 ; TITLE OF INVENTION: SOLID SUPPORTS WITH MINICELLS  
 ; FILE REFERENCE: MPEX.008DV4  
 ; CURRENT APPLICATION NUMBER: US/10/157,166  
 ; CURRENT FILING DATE: 2002-05-28  
 ; PRIOR APPLICATION NUMBER: 60/293,566  
 ; PRIOR FILING DATE: 2001-05-24  
 ; PRIOR APPLICATION NUMBER: 60/359,843  
 ; PRIOR FILING DATE: 2002-02-25  
 ; PRIOR APPLICATION NUMBER: 10/154,951  
 ; PRIOR FILING DATE: 2002-05-24  
 ; NUMBER OF SEQ ID NOS: 257  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 186  
 ; LENGTH: 3284  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Fusion protein  
 ; US-10-157-166-186

Query Match 36.9%; Score 28.4; DB 14; Length 3284;  
 Best Local Similarity 96.7%; Pred. No. 0.94;  
 Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTGCAT 52  
 |||||  
 Db 1254 CGAGCATCATCCTCTGCATGGTCAGGTGCAT 1283

## RESULT 31

US-10-156-902-186

; Sequence 186, Application US/10156902  
 ; Publication No. US20030199088A1

GENERAL INFORMATION:

APPLICANT: Sabbadini, Roger A.

APPLICANT: Neil Berkley

APPLICANT: Surber, Mark W.

TITLE OF INVENTION: MINICELL-BASED GENE THERAPY

FILE REFERENCE: MPEX.008DV16

CURRENT APPLICATION NUMBER: US/10/156,902

CURRENT FILING DATE: 2002-05-28

PRIOR APPLICATION NUMBER: 60/293,566

PRIOR FILING DATE: 2001-05-24

PRIOR APPLICATION NUMBER: 60/359,843

PRIOR FILING DATE: 2002-02-25

PRIOR APPLICATION NUMBER: 10/154,951

PRIOR FILING DATE: 2002-05-24

NUMBER OF SEQ ID NOS: 257

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 186

LENGTH: 3284

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Fusion protein

US-10-156-902-186

Query Match 36.9%; Score 28.4; DB 14; Length 3284;  
 Best Local Similarity 96.7%; Pred. No. 0.94;  
 Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTGCAT 52  
 |||||  
 Db 1254 CGAGCATCATCCTCTGCATGGTCAGGTGCAT 1283

## RESULT 32

US-10-157-318-186

; Sequence 186, Application US/10157318  
 ; Publication No. US20030199089A1

GENERAL INFORMATION:

APPLICANT: Surber, Mark W.



Sabbadini, Roger A.  
ENTION: MEMBRANE TO MEMBRANE DELIVERY  
CE: MPEX.008DV8  
ICATION NUMBER: US/10/157,318  
NG DATE: 2002-05-28  
ACTION NUMBER: 60/293,566  
DATE: 2001-05-24  
ATION NUMBER: 60/359,843  
DATE: 2002-02-25  
ATION NUMBER: 10/154,51  
DATE: 2002-05-24  
Q ID NOS: 257  
stSEQ for Windows Version 4.0

4

rtificial Sequence

MATION: Fusion protein

86

36.9%; Score 28.4; DB 14; Length 3284;  
milarity 96.7%; Pred. No. 0.94;  
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGTCATGGTCAGGTCAT 52  
|||||  
GAGCATCATCCTCTGTCATGGTCAGGTCAT 1283

86

Application US/10156811  
o. US20030207833A1  
MATION:  
eil Berkley  
Robert Klepper  
Sabbadini, Roger A.  
ENTION: PHARMACEUTICAL COMPOSITIONS WITH  
ENTION: MINICELLS  
CE: MPEX.008DV9  
ICATION NUMBER: US/10/156,811  
NG DATE: 2002-05-28  
ACTION NUMBER: 60/293,566  
DATE: 2001-05-24  
ATION NUMBER: 60/359,843  
DATE: 2002-02-25  
ATION NUMBER: 10/154,951  
DATE: 2002-05-24  
Q ID NOS: 257  
stSEQ for Windows Version 4.0

4

rtificial Sequence

MATION: Fusion protein

86

36.9%; Score 28.4; DB 15; Length 3284;  
milarity 96.7%; Pred. No. 0.94;  
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGTCATGGTCAGGTCAT 52  
|||||  
GAGCATCATCCTCTGTCATGGTCAGGTCAT 1283

186

Application US/10157320A  
o. US20030219408A1  
MATION:

; APPLICANT: Sabbadini, Roger A.  
; APPLICANT: Klepper, Robert  
; TITLE OF INVENTION: METHODS OF MAKING PHARMACEUTICAL  
; FILE REFERENCE: MPEX.008DV10  
; CURRENT APPLICATION NUMBER: US/10/157,320A  
; CURRENT FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: 10/154,951  
; PRIOR FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/359,843  
; PRIOR FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 60/293,566  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 258  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 186  
; LENGTH: 3284  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Gene encoding a fusion protein  
US-10-157-320A-186

Query Match 36.9%; Score 28.4; DB 15; Length 3284;  
Best Local Similarity 96.7%; Pred. No. 0.94;  
Matches 29; Conservative 0; Mismatches 1; Indels 0;

QY 23 CTAGCATCATCCTCTGTCATGGTCAGGTCAT 52  
|||||  
Db 1254 CGAGCATCATCCTCTGTCATGGTCAGGTCAT 1283

RESULT 35

US-10-157-418A-186  
; Sequence 186, Application US/10157418A  
; Publication No. US20030219888A1  
; GENERAL INFORMATION:  
; APPLICANT: Segall, Anca M.  
; APPLICANT: Klepper, Robert  
; TITLE OF INVENTION: MINICELL-BASED BIOREMEDIATION  
; FILE REFERENCE: MPEX.008DV22  
; CURRENT APPLICATION NUMBER: US/10/157,418A  
; CURRENT FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: 10/154,951  
; PRIOR FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/359,843  
; PRIOR FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 60/293,566  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 258  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 186  
; LENGTH: 3284  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Gene encoding a fusion protein  
US-10-157-418A-186

Query Match 36.9%; Score 28.4; DB 15; Length 3284;  
Best Local Similarity 96.7%; Pred. No. 0.94;  
Matches 29; Conservative 0; Mismatches 1; Indels 0;

QY 23 CTAGCATCATCCTCTGTCATGGTCAGGTCAT 52  
|||||  
Db 1254 CGAGCATCATCCTCTGTCATGGTCAGGTCAT 1283

RESULT 36

US-10-157-317-186  
; Sequence 186, Application US/10157317  
; Publication No. US2003023335A1  
; GENERAL INFORMATION:

Surber, Mark W.  
 Sabbadini, Roger A.  
 Neil Berkley  
 /ENTION: MINICELL-BASED SCREENING FOR COMPOUNDS  
 /ENTION: AND PROTEINS THAT MODULATE THE ACTIVITY OF SIGNALLING  
 /ENTION: PROTEINS  
 /CE: MPEX.008DV21

/ICATION NUMBER: US/10/157,317  
 /NG DATE: 2002-05-28  
 /ATION NUMBER: 60/293,566  
 / DATE: 2001-05-24  
 /ATION NUMBER: 60/359,843  
 / DATE: 2002-02-25  
 /ATION NUMBER: 10/154,951  
 / DATE: 2002-05-24  
 /Q ID NOS: 257  
 /stSEQ for Windows Version 4.0

14

Artificial Sequence

/MATION: Fusion protein

.86

36.9%; Score 28.4; DB 15; Length 3284;  
 ilarity 96.7%; Pred. No. 0.94; 1; Indels 0; Gaps 0;  
 . Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52  
 |||||  
 XGAGCATCATCCTCTGCATGGTCAGGTCAT 1283

.86

Application US/10157339  
 /o. US200400057000A1

/MATION:

Surber, Mark W.

Matthew Giacalone

/ENTION: POROPLASTS

/CE: MPEX.008DV3

/ICATION NUMBER: US/10/157,339

/NG DATE: 2002-08-29

/ATION NUMBER: 60/293,566

/ DATE: 2001-05-24

/ATION NUMBER: 60/359,843

/ DATE: 2002-02-25

/ATION NUMBER: 10/154,951

/ DATE: 2002-05-24

/Q ID NOS: 257

/stSEQ for Windows Version 4.0

14

Artificial Sequence

/MATION: Fusion protein

.186

36.9%; Score 28.4; DB 15; Length 3284;  
 ilarity 96.7%; Pred. No. 0.94;  
 ; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52  
 |||||  
 XGAGCATCATCCTCTGCATGGTCAGGTCAT 1283

32

Application US/10338411

; Publication No. US20030153045A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Butt, Tauseef  
 ; APPLICANT: Weeks, Stephen  
 ; APPLICANT: Tran, Hiep  
 ; APPLICANT: Malakhova, Oxana  
 ; APPLICANT: Malakhov, Micheal  
 ; TITLE OF INVENTION: Methods and Compositions for Protein  
 ; TITLE OF INVENTION: Expression and Purification  
 ; FILE REFERENCE: 1955-2792US1  
 ; CURRENT APPLICATION NUMBER: US/10/338,411  
 ; CURRENT FILING DATE: 2003-01-07  
 ; PRIOR APPLICATION NUMBER: 60/346,449  
 ; PRIOR FILING DATE: 2002-01-07  
 ; NUMBER OF SEQ ID NOS: 65  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO 32  
 ; LENGTH: 3396  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetic Sequence  
 US-10-338-411-32

Query Match 36.9%; Score 28.4; DB 14; Length 3396.  
 Best Local Similarity 96.7%; Pred. No. 0.95; 1; Indels 0.  
 Matches 29; Conservative 0; Mismatches 1; Indels 0.

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52  
 |||||  
 Db 1392 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1421

RESULT 39

US-10-389-640-32  
 ; Sequence 32, Application US/10389640

; Publication No. US20040018591A1  
 ; GENERAL INFORMATION:

; APPLICANT: Butt, Tauseef  
 ; APPLICANT: Weeks, Stephen

; APPLICANT: Tran, Hiep  
 ; APPLICANT: Malakhova, Oxana

; APPLICANT: Malakhov, Micheal  
 ; TITLE OF INVENTION: Methods and Compositions for Protein

; TITLE OF INVENTION: Expression and Purification  
 ; FILE REFERENCE: 1955-P02972US2

; CURRENT APPLICATION NUMBER: US/10/389,640  
 ; CURRENT FILING DATE: 2003-03-14

; PRIOR APPLICATION NUMBER: 10/338,411  
 ; PRIOR FILING DATE: 2003-01-07

; PRIOR APPLICATION NUMBER: 60/346,449  
 ; PRIOR FILING DATE: 2002-01-07

; NUMBER OF SEQ ID NOS: 65  
 ; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 32  
 ; LENGTH: 3396

; TYPE: DNA  
 ; ORGANISM: Artificial Sequence

; FEATURE:  
 ; OTHER INFORMATION: Synthetic Sequence

US-10-389-640-32

Query Match 36.9%; Score 28.4; DB 15; Length 3396;  
 Best Local Similarity 96.7%; Pred. No. 0.95; 1; Indels 0;  
 Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52  
 |||||  
 Db 1392 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1421

RESULT 40

US-09-956-998A-1

publication US/09956998A  
2002082236A1  
MARION:  
lack Jr., Charles A.  
ENTION: COMPOSITIONS AND METHODS FOR ACTIVATING  
ENTION: GENES OF INTEREST  
CE: 5722-2(35722/191928)  
ICATION NUMBER: US/09/956,998A  
NG DATE: 2001-09-20  
ATION NUMBER: 09/446,402  
DATE: 1999-12-20  
ATION NUMBER: 60/050,772  
DATE: 1997-06-25  
Q ID NOS: 19  
stSEQ for Windows Version 4.0  
9  
rtificial Sequence  
MARION: Recombinant Molecule containing multiple cloning  
MARION: site, kozak sequence, lacZ gene.  
isc\_feature  
1)...(64)  
MARION: Multiple cloning site  
isc\_feature  
65)...(79)  
MARION: Consensus sequence for the "Kozak sequence"  
MARION: (translation initiation)  
rim transcript  
80)...(4279)  
MARION: Beta galactosidase  
1  
milarity 36.8%; Score 28.4; DB 9; Length 4279;  
Conservative 96.7%; Pred. No. 1;  
Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
TAGCATCATCCTCTGCATCGTCAGGTCAAT 52  
|||||  
SAGCATCATCCTCTGCATCGTCAGGTCAAT 1864  
3: April 13, 2004, 21:46:12  
5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

ucleic search, using sw model

April 13, 2004, 18:19:17 ; Search time 43 Seconds  
(without alignments)  
993.749 Million cell updates/sec

US-10-090-326-25

77

1 actaaccagccagcanca.....99agnacacagcancagt 77

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

682709 seqs, 277475446 residues

hits satisfying chosen parameters: 1365418

length: 0

length: 2000000000

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6A COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/6B COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/PCTUS COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

is the number of results predicted by chance to have a  
alter than or equal to the score of the result being printed,  
rived by analysis of the total score distribution.

#### SUMMARIES

Query Match	Length	DB	ID	Description
39.0	1059	2	US-08-484-956-58	Sequence 58, Appl
39.0	1059	2	US-08-484-956-59	Sequence 59, Appl
39.0	1059	2	US-08-757-653-58	Sequence 58, Appl
39.0	1059	2	US-08-757-653-59	Sequence 59, Appl
39.0	1059	4	US-08-520-946-58	Sequence 58, Appl
39.0	1059	4	US-08-520-946-59	Sequence 59, Appl
39.0	1059	4	US-09-655-378A-58	Sequence 58, Appl
39.0	1059	4	US-09-655-378A-59	Sequence 59, Appl
39.0	1587	2	US-08-484-956-60	Sequence 60, Appl
39.0	1587	2	US-08-484-956-61	Sequence 61, Appl
39.0	1587	2	US-08-757-653-60	Sequence 60, Appl
39.0	1587	2	US-08-757-653-61	Sequence 61, Appl
39.0	1587	4	US-08-520-946-60	Sequence 60, Appl
39.0	1587	4	US-08-520-946-61	Sequence 61, Appl
39.0	1587	4	US-09-655-378A-60	Sequence 60, Appl
39.0	1587	4	US-09-655-378A-61	Sequence 61, Appl
39.0	1894	1	US-08-233-305-1	Sequence 1, Appl
39.0	1894	1	US-08-203-054-1	Sequence 1, Appl
39.0	1894	1	US-08-636-676-1	Sequence 1, Appl
39.0	1894	2	US-08-545-212-1	Sequence 1, Appl
39.0	1894	2	US-08-583-238-1	Sequence 1, Appl
39.0	1905	3	US-08-540-922D-1	Sequence 1, Appl
39.0	1906	1	US-08-081-673-1	Sequence 1, Appl
39.0	1906	1	US-08-587-391-1	Sequence 1, Appl
39.0	1906	4	US-10-011-436-1	Sequence 1, Appl
39.0	1910	2	US-08-370-909-18	Sequence 18, Appl
39.0	2384	1	US-08-181-471-1	Sequence 1, Appl

28	30	39.0	2384	3	US-09-056-105-3	Sequenc
29	28.4	36.9	3096	4	US-09-489-039A-4015	Sequenc
30	28.4	36.9	3152	1	US-07-924-028A-3	Sequenc
31	28.4	36.9	3365	1	US-07-789-915A-1	Sequenc
32	28.4	36.9	3365	1	US-08-005-002C-1	Sequenc
33	28.4	36.9	3365	1	US-08-487-203A-1	Sequenc
34	28.4	36.9	4279	4	US-09-446-402A-1	Sequenc
35	28.4	36.9	4810	3	US-08-852-629-11	Sequenc
36	28.4	36.9	4838	3	US-08-852-629-15	Sequenc
37	28.4	36.9	6700	4	US-09-654-449-1	Sequenc
38	28.4	36.9	7252	4	US-09-238-356-27	Sequenc
39	28.4	36.9	7366	6	5169760-3	Patent 1
40	28.4	36.9	7387	4	US-09-238-356-28	Sequenc
41	28.4	36.9	7408	4	US-09-115-475-13	Sequenc
42	28.4	36.9	7612	4	US-09-700-934A-2	Sequenc
43	28.4	36.9	7664	4	US-08-994-689C-10	Sequenc
44	28.4	36.9	7852	3	US-08-836-022A-2	Sequenc
45	28.4	36.9	7852	3	US-09-427-048A-2	Sequenc

#### ALIGNMENTS

RESULT 1  
US-08-484-956-58  
; Sequence 58, Application US/08484956  
; Patent No. 5843654  
; GENERAL INFORMATION:  
; APPLICANT: DAHLBERG, JAMES E.  
; APPLICANT: LYAMICHEV, VICTOR I.  
; APPLICANT: BROW, MARY ANN D.  
; APPLICANT: OLDENBURG, MARY C.  
; APPLICANT: HEISLER, LAURA  
; TITLE OF INVENTION: DETECTION OF p53 MUTATIONS  
; NUMBER OF SEQUENCES: 114  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HAVERSTOCK, MEDLEN & CARROLL  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,956  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/402,601  
; FILING DATE: 09-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/337,164  
; FILING DATE: 09-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/254,359  
; FILING DATE: 06-JUN-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/073,384  
; FILING DATE: 04-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/986,330  
; FILING DATE: 07-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CARPOLL, PETER G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: FORS-01801  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338

: (415) 397-8338

60 AGNACACAGCCANCACT 77



REGISTRATION INFORMATION: FORS-01756

TELEPHONE: (415) 705-8410

FOR SEQ ID NO: 59:

CHARACTERISTICS:

1059 base pairs

nucleic acid

NESS: single

TYPE: linear

TYPE: DNA (genomic)

9

Similarity 39.0%; Score 30; DB 4; Length 1059;

Similarity 67.9%; Pred. No. 0.026; 24; Indels 1; Gaps 1;

Conservative 0; Mismatches 24; Indels 1; Gaps 1;

ACTTACNAGCCAGCCAGCAGCATCTTCTCTGTCATG-GTCAGGTGATTGGAGG 59

ACTTACTAGCCAGCCAGCAGCATCTTCTCTGTCATG-GTCAGGTGATTGGAGG 534

AGNACACAGCCAGCAGT 77

AGTACACAGCCAGCAGT 552

58

Application US/09655378A

73616

FORMATION:

INVENTOR: BROW, MARY ANN D.

LYAMICHEV, VICTOR I.

OLIVE, DAVID M.

OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF

PATHOGENS

OF SEQUENCES: 165

ADDRESS: MEDLEN & CARROLL

STREET: 220 MONTGOMERY STREET, SUITE 2200

CITY: SAN FRANCISCO

STATE: CALIFORNIA

COUNTRY: UNITED STATES OF AMERICA

ZIP: 94104

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

AT APPLICATION DATA:

APPLICATION NUMBER: US/09/655,378A

FILING DATE: 05-Sep-2000

CLASSIFICATION: <Unknown>

KEY/AGENT INFORMATION:

NAME: CARROLL, PETER G.

REGISTRATION NUMBER: 32,837

REFERENCE/DOCKET NUMBER: FORS-01756

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

FOR SEQ ID NO: 58:

CHARACTERISTICS:

LENGTH: 1059 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

FILE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 58:

58

Similarity 39.0%; Score 30; DB 4; Length 1059;

Similarity 67.9%; Pred. No. 0.026; 24; Indels 1; Gaps 1;

Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 1 ACTTACNAGCCAGCCAGCAGCATCTTCTCTGTCATG-GTCAGGTGATTGGAGG 59

Db 475 ACTTACTAGCCAGCCAGCAGCATCTTCTCTGTCATG-GTCAGGTGATTGGAGG 534

QY 60 AGNACACAGCCAGCAGT 77

Db 535 AGTACACAGCCAGCAGT 552

RESULT 8

US-09-655-378A-59

; Sequence 59, Application US/09655378A

; Patent No. 6673616

; GENERAL INFORMATION:

; APPLICANT: BROW, MARY ANN D.

; LYAMICHEV, VICTOR I.

; OLIVE, DAVID M.

; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF

; PATHOGENS

; NUMBER OF SEQUENCES: 165

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MEDLEN & CARROLL

; STREET: 220 MONTGOMERY STREET, SUITE 2200

; STATE: CALIFORNIA

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/655,378A

; FILING DATE: 05-Sep-2000

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: CARROLL, PETER G.

; REGISTRATION NUMBER: 32,837

; REFERENCE/DOCKET NUMBER: FORS-01756

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 705-8410

; TELEFAX: (415) 397-8338

; INFORMATION FOR SEQ ID NO: 59:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1059 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; SEQUENCE DESCRIPTION: SEQ ID NO: 59:

US-09-655-378A-59

Query Match 39.0%; Score 30; DB 4; Length 1059;

Best Local Similarity 67.9%; Pred. No. 0.026; 24; Indels 1;

Matches 53; Conservative 0; Mismatches 24; Indels 1;

QY 1 ACTTACNAGCCAGCCAGCAGCATCTTCTCTGTCATG-GTCAGGTGATTGGAGG 59

Db 475 ACTTACTAGCCAGCCAGCAGCATCTTCTCTGTCATG-GTCAGGTGATTGGAGG 534

QY 60 AGNACACAGCCAGCAGT 77

Db 535 AGTACACAGCCAGCAGT 552

RESULT 9

US-08-484-956-60

; Sequence 60, Application US/08484956

; Patent No. 5843654

; GENERAL INFORMATION:

; APPLICANT: DAHLBERG, JAMES E.

APPLICANT: BROW, MARY ANN D.  
 APPLICANT: OLDENBURG, MARY C.  
 APPLICANT: HEISLER, LAURA  
 TITLE OF INVENTION: DETECTION OF P53 MUTATIONS  
 NUMBER OF SEQUENCES: 114  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: HAYESTOCK, MEDLEN & CARROLL  
 STREET: 220 MONTGOMERY STREET, SUITE 2200  
 CITY: SAN FRANCISCO  
 STATE: CALIFORNIA  
 COUNTRY: UNITED STATES OF AMERICA  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/484,956  
 FILING DATE:

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/402,601  
FILING DATE: 09-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/337,164  
FILING DATE: 09-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/254,359  
FILING DATE: 06-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/073,384  
FILING DATE: 04-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/986,330  
FILING DATE: 07-DEC-1992  
NAME: CARROLL J, PETER G.  
REGISTRATION NUMBER: 32,1837  
REFERENCE/DOCKET NUMBER: FORS-01801  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1587 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
484-956-61

[illegible]

RESULT 11  
US-08-757-653-60  
; Sequence 60, Application US/08757653  
; Patent No. 5843669  
; GENERAL INFORMATION:  
; APPLICANT: Kaiser, Michael W.  
; APPLICANT: Lymanichev, Victor I.  
; APPLICANT: Lymanichev, Victor I.



READABLE FORM:  
TYPE: Floppy disk  
R: IBM PC compatible  
NG SYSTEM: PC-DOS/MS-DOS  
E: Patentin Release #1.0, Version #1.30  
APPLICATION DATA:  
TION NUMBER: US/08/757.653

```

/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1587 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
US-08-520-946-60

```

39.0%; Score 30; DB 4; Length 1587;  
milarity 67.9%; Pred.No. 0.03;  
Conservative 0; Mismatches 24; Indels 1; Gaps 1;  
CTTACNAGCCGACGANCATTCTAGCATCATCTCTGCATG-GTCAGGTCAITTTGAGG 59  
CTTACTCAGCCGACGATCATTTCTCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841  
GNACACAGCCGANCAGT 77  
GTTCACAGCCATCATG 859

1 Application US/08520946  
72424  
INFORMATION:  
BROW, MARY ANN D.  
LYANICHEV, VICTOR I.  
OLIVE, DAVID M.  
INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
AGENTS: PATHOGENS  
SEQUENCES: 160  
SEQUENCE ADDRESS:  
E: MEDLEN & CARROLL  
220 MONTGOMERY STREET, SUITE 2200  
SAN FRANCISCO  
CALIFORNIA  
UNITED STATES OF AMERICA  
104

```

READABLE FORM:
#P: Floppy disk
# IBM PC compatible
# SYSTEM: PC-DOS/MS-DOS
: PatentIn Release #1.0, Version #1.25
APPLICATION DATA:
[CON NUMBER: US/08/520.946
```

DATE: 4/25/68  
 LOCATION: 435  
 IDENTIFICATION:  
 ARROLL, PETER G.  
 SECTION NUMBER: 32,837  
 EX/DOCKET NUMBER: FORS-01756

```

[CATION INFORMATION
3: (415) 705-8410
3: (415) 397-8338
FOR SEQ ID NO: 61:
CHARACTERISTICS:
1587 base pairs
nucleic acid
NESS: single
: linear
:PE: DNA (genomic)

```

```

39.0%; Score 30; DB 4; Length 1587;
ilarity 67.9%; Pred. No. 0.03;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

TTTACNAGCCCGAGCAGCATCTCTAGCATCTCCCTCTGCGATG-GTCAGGTCATTGTGAGG 59
TTTACTCAGCCCGAGCATCATCTCTCCCTCTTGGCAGATTGCTGTAGCCGATTGGAGG 841

NNACACAGCCCGAGCAT 77
TTACACAGCCCATCAGT 859

```

50  
Application US/09655378A

```

; Patent No. 6673616
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; LYAMICHEV, VICTOR I.
; OLIVE, DAVID M.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; PATHOGENS
;
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/655,378A
; FILING DATE: 05-Sep-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
;
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
;
; INFORMATION FOR SEQ ID NO: 60:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1587 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-655-378A-60

```

Query Match	39.0%	Score 30;	DB 4;	Length 1587;
Best Local Similarity	67.9%	Pred. No. 0.03;		
Matches	53;	Conservative	0;	Mismatches 24; Indels 1;
QY	1	ACTTACNCGCCGAGCANCATTTAGCATCATCTCTGCGATG-GTCAGGTCAT		
Db	782	ACTTACTCAGCCGAGCATCTTCTCTTGGCAGATTGTCTGTAGCCGA		
QY	60	AGNACAACGCCANCAGT	77	
Db	842	AGTACAACGCCATCAGT	859	

RESULT 16  
 US-09-655-378A-61  
 ? Sequence 61, Application US/09655378A  
 ? Patent No. 6673616  
 ?  
 ? GENERAL INFORMATION:  
 ?  
 ? APPLICANT: BROW, MARY ANN D.  
 ? LYAMICHEV, VICTOR I.  
 ? OLIVE, DAVID M.  
 ?  
 ? TITLE OF INVENTION: DETECTION AND IDENTIFICATION OF  
 ? PATHOGENS  
 ?  
 ? NUMBER OF SEQUENCES: 165  
 ? CORRESPONDENCE ADDRESS:  
 ? ADDRESS: MEDLEN & CARROLL  
 ? STREET: 220 MONTGOMERY STREET, SUITE 2200  
 ? CITY: SAN FRANCISCO  
 ? STATE: CALIFORNIA  
 ? COUNTRY: UNITED STATES OF AMERICA  
 ? ZIP: 94104  
 ?  
 ? COMPUTER READABLE FORM:

; APPLICATION NUMBER: 07/994,928  
 ; FILING DATE: 22-DEC-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Hanson, No. 5519117man D.  
 ; REGISTRATION NUMBER: 30,946  
 ; REFERENCE/DOCKET NUMBER: LUD 5360.1  
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1894 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-233-305-1

Query Match	39.0%	Score 30;
Best Local Similarity	67.9%	Pred. No.
Matches	53;	Conservative 0; Mismatch

QY 1 ACTTACACGCCGACCATCTTAGCAT

DG		60	AGNACAACAGCCATCAGT	77
/89	ACTTA	CAGCC	CAGCATCACTTC	TCTCTC
Dd		848	AGTACAACAGCCATCAGT	865

RESULT 18  
US-08-203-054-1

```

: Sequence 1, Application US/08203054
: Patent No. 553096
:
: GENERAL INFORMATION:
:
: APPLICANT: Wlfel, Thomas; Van Pel, A
:
: APPLICANT: Boon-Falleur, Thierry
:
: TITLE OF INVENTION: ISOLATED, TYROSIN
:
: TITLE OF INVENTION: USES THEREOF
:
: NUMBER OF SEQUENCES: 5
:
: CORRESPONDENCE ADDRESS:

```

ADDRESS: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360K  
COMPUTER: IBM PS/2

```

/ OPERATING SYSTEM: PC-DOS
/ SOFTWARE: Wordperfect
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/203,054
/ FILING DATE: 28-FEB-1994
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/081,673
/ FILING DATE: 23-JUN-1990

```

7 FILING DATE: 23-JUNE-1993  
 7 PRIOR APPLICATION DATA:  
 7 APPLICATION NUMBER: 08/005,714  
 7 FILING DATE: 28-APRIL-1993  
 7 PRIOR APPLICATION DATA:  
 7 APPLICATION NUMBER: 07/994,928  
 7 FILING DATE: 22-DEC-1992  
 7 ATTORNEY/AGENT INFORMATION:  
 7 NAME: Hansson, NO. 5530096man D.  
 7

REGISTRATION NUMBER: 30,946  
REFERENCE/POCKET NUMBER: IUD 360  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 638-3864  
INFORMATION FOR SEQ ID NO: 1:



```

/ Patent No. 6284476
/
/ GENERAL INFORMATION:
/
/ APPLICANT: Boon-Falleur, Thierry; Brichard, Vincent; Van
/ APPLICANT: Pel, Aline; De Plaen, Etienne; Coulie, Pierre;
/ APPLICANT: Renaud Jean-Christophe; Wolfel, Thomas; and
/ APPLICANT: Leche, Bernard.
/
/ TITLE OF INVENTION: METHOD OF IDENTIFYING INDIVIDUALS SUFFER
/ TITLE OF INVENTION: FROM A CELLULAR ABNORMALITY SOME OF WHOM
/ TITLE OF INVENTION: ABNORMAL CELLS PRESENT COMPLEXES OF HUMA
/ TITLE OF INVENTION: LEUKOCYTE ANTIGEN TYROSINASE DERIVED
/ TITLE OF INVENTION: PEPTIDES, AND METHODS FOR TREATING SAID
/ TITLE OF INVENTION: INDIVIDUALS
/
/ NUMBER OF SEQUENCES: 12
/
/ CORRESPONDENCE ADDRESS:
/
/ ADDRESSEE: Felfe & Lynch
/ STREET: 805 Third Avenue
/ CITY: New York
/ STATE: New York
/ COUNTRY: USA
/
/ ZIP: 10022
/
/ COMPUTER READABLE FORM:
/
/ MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
/
/ OPERATING SYSTEM: PC-DOS
/ SOFTWARE: Wordperfect
/
/ CURRENT APPLICATION DATA:
/
/ APPLICATION NUMBER: US/08/540,922D
/
/ FILING DATE: October 11, 1995
/
/ CLASSIFICATION: 514
/
/ PRIORITY APPLICATION DATA:
/
/ APPLICATION NUMBER: 08/054,714
/
/ FILING DATE: 28 April 1993
/
/ PRIOR APPLICATION DATA:
/
/ APPLICATION NUMBER: 07/994,928
/
/ FILING DATE: 22 December 1992
/
/ ATTORNEY/AGENT INFORMATION:
/
/ NAME: Mary Anne Schofield
/
/ REGISTRATION NUMBER: 36,669
/
/ REFERENCE/DOCKET NUMBER: LUD 5299.5
/
/ TELECOMMUNICATION INFORMATION:
/
/ TELEPHONE: (212) 688-9200
/
/ TELEFAX: (212) 838-3884
/
/ INFORMATION FOR SEQ ID NO: 1:
/
/ SEQUENCE CHARACTERISTICS:
/
/ LENGTH: 1905 nucleotides
/
/ TYPE: nucleic acid
/
/ STRANDEDNESS: single
/
/ TOPOLOGY: linear
/
/ US-08-540-922D-1

```

```

Query Match      39.0%; Score 30; DB 3; Length 1905;
Best Local Similarity 67.9%; Pred. No. 0.032;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy      1  ACTTACNAGCCGACGANCATTCTAGCATCATCTCTGCAAG-GTCAGGTCA
Db      806  ACTTACTCAGCCGAGCATCTCTCTCTTTGGCGATTGTCTGTAGCGG
Qy      60  AGNACAACAGCCNACAGT 77
Db      865  AGTACAACAGCCATCAGT 883

```

RESULT 23  
US-08-081-673-1  
; Sequence 1, Application US/08081673  
; Patent No. 5487974  
; GENERAL INFORMATION:  
; APPLICANT: Boon-Falleur, Thierry; Brichard, Vincent; Van Pel,  
; APPLICANT: Aline; Daplaen, Etienne; Coulle, Pierre; Renaud,  
; APPLICANT: Wolfel, Thomas; Lethe, Bernard  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING INDIVIDUALS SUFFER  
; TITLE OF INVENTION: FROM A CELLULAR ABNORMALITY SOME OF WHOSE

TYPE: Diskette, 5.25 inch, 360 kb storage

JG SYSTEM: PC-DOS

Y: NOT PERFECT  
REPLICATION DATA.

ATION NUMBER: US/08

[CATION: 435]

ATION NUMBER: 08/200

DATE: 28-FEB-1994  
-CATION. 435

NOTATION NUMBER: 08/08

LOCATION DATA:

DATE: 28-APRIL-1993

ICATION DATA: 07/99

DATE: 22-DEC-1992  
CENT: TROOPMENT

lanson, No. 5843688

FILE/DOCKET NUMBER:

TELEPHONE: (212) 688-9200

(212) 838-3884  
FOR CEO TO NO. 1

### CHARACTERISTICS:

nucleic acid<sup>2</sup> $\gamma$ : linear

39.08;

Conservative

ACTTACNCAGCCCAGCAGC

CTTACTACCCACCATC

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

1. *Introduction*  
 2. *Methodology*  
 3. *Results*  
 4. *Discussion*  
 5. *Conclusion*  
 6. *References*  
 7. *Appendix*  
 8. *Notes*  
 9. *Tables*  
 10. *Figures*  
 11. *Supplementary Materials*  
 12. *Index*  
 13. *Abstract*  
 14. *Keywords*  
 15. *Subject Headings*  
 16. *Summary*  
 17. *References*  
 18. *Appendix*  
 19. *Notes*  
 20. *Tables*  
 21. *Figures*  
 22. *Supplementary Materials*  
 23. *Index*  
 24. *Abstract*  
 25. *Keywords*  
 26. *Subject Headings*  
 27. *Summary*  
 28. *References*  
 29. *Appendix*  
 30. *Notes*  
 31. *Tables*  
 32. *Figures*  
 33. *Supplementary Materials*  
 34. *Index*  
 35. *Abstract*  
 36. *Keywords*  
 37. *Subject Headings*  
 38. *Summary*  
 39. *References*  
 40. *Appendix*  
 41. *Notes*  
 42. *Tables*  
 43. *Figures*  
 44. *Supplementary Materials*  
 45. *Index*  
 46. *Abstract*  
 47. *Keywords*  
 48. *Subject Headings*  
 49. *Summary*  
 50. *References*  
 51. *Appendix*  
 52. *Notes*  
 53. *Tables*  
 54. *Figures*  
 55. *Supplementary Materials*  
 56. *Index*  
 57. *Abstract*  
 58. *Keywords*  
 59. *Subject Headings*  
 60. *Summary*  
 61. *References*  
 62. *Appendix*  
 63. *Notes*  
 64. *Tables*  
 65. *Figures*  
 66. *Supplementary Materials*  
 67. *Index*  
 68. *Abstract*  
 69. *Keywords*  
 70. *Subject Headings*  
 71. *Summary*  
 72. *References*  
 73. *Appendix*  
 74. *Notes*  
 75. *Tables*  
 76. *Figures*  
 77. *Supplementary Materials*  
 78. *Index*  
 79. *Abstract*  
 80. *Keywords*  
 81. *Subject Headings*  
 82. *Summary*  
 83. *References*  
 84. *Appendix*  
 85. *Notes*  
 86. *Tables*  
 87. *Figures*  
 88. *Supplementary Materials*  
 89. *Index*  
 90. *Abstract*  
 91. *Keywords*  
 92. *Subject Headings*  
 93. *Summary*  
 94. *References*  
 95. *Appendix*  
 96. *Notes*  
 97. *Tables*  
 98. *Figures*  
 99. *Supplementary Materials*  
 100. *Index*  
 101. *Abstract*  
 102. *Keywords*  
 103. *Subject Headings*  
 104. *Summary*  
 105. *References*  
 106. *Appendix*  
 107. *Notes*  
 108. *Tables*  
 109. *Figures*  
 110. *Supplementary Materials*  
 111. *Index*  
 112. *Abstract*  
 113. *Keywords*  
 114. *Subject Headings*  
 115. *Summary*  
 116. *References*  
 117. *Appendix*  
 118. *Notes*  
 119. *Tables*  
 120. *Figures*  
 121. *Supplementary Materials*  
 122. *Index*  
 123. *Abstract*  
 124. *Keywords*  
 125. *Subject Headings*  
 126. *Summary*  
 127. *References*  
 128. *Appendix*  
 129. *Notes*  
 130. *Tables*  
 131. *Figures*  
 132. *Supplementary Materials*  
 133. *Index*  
 134. *Abstract*  
 135. *Keywords*  
 136. *Subject Headings*  
 137. *Summary*  
 138. *References*  
 139. *Appendix*  
 140. *Notes*  
 141. *Tables*  
 142. *Figures*  
 143. *Supplementary Materials*  
 144. *Index*  
 145. *Abstract*  
 146. *Keywords*  
 147. *Subject Headings*  
 148. *Summary*  
 149. *References*  
 150. *Appendix*  
 151. *Notes*  
 152. *Tables*  
 153. *Figures*  
 154. *Supplementary Materials*  
 155. *Index*  
 156. *Abstract*  
 157. *Keywords*  
 158. *Subject Headings*  
 159. *Summary*  
 160. *References*  
 161. *Appendix*  
 162. *Notes*  
 163. *Tables*  
 164. *Figures*  
 165. *Supplementary Materials*  
 166. *Index*  
 167. *Abstract*  
 168. *Keywords*  
 169. *Subject Headings*  
 170. *Summary*  
 171. *References*  
 172. *Appendix*  
 173. *Notes*  
 174. *Tables*  
 175. *Figures*  
 176. *Supplementary Materials*  
 177. *Index*  
 178. *Abstract*  
 179. *Keywords*  
 180. *Subject Headings*  
 181. *Summary*  
 182. *References*  
 183. *Appendix*  
 184. *Notes*  
 185. *Tables*  
 186. *Figures*  
 187. *Supplementary Materials*  
 188. *Index*  
 189. *Abstract*  
 190. *Keywords*  
 191. *Subject Headings*  
 192. *Summary*  
 193. *References*  
 194. *Appendix*  
 195. *Notes*  
 196. *Tables*  
 197. *Figures*  
 198. *Supplementary Materials*  
 199. *Index*  
 200. *Abstract*  
 201. *Keywords*  
 202. *Subject Headings*  
 203. *Summary*  
 204. *References*  
 205. *Appendix*  
 206. *Notes*  
 207. *Tables*  
 208. *Figures*  
 209. *Supplementary Materials*  
 210. *Index*  
 211. *Abstract*  
 212. *Keywords*  
 213. *Subject Headings*  
 214. *Summary*  
 215. *References*  
 216. *Appendix*  
 217. *Notes*  
 218. *Tables*  
 219. *Figures*  
 220. *Supplementary Materials*  
 221. *Index*  
 222. *Abstract*  
 223. *Keywords*  
 224. *Subject Headings*  
 225. *Summary*  
 226. *References*  
 227. *Appendix*  
 228. *Notes*  
 229. *Tables*  
 230. *Figures*  
 231. *Supplementary Materials*  
 232. *Index*  
 233. *Abstract*  
 234. *Keywords*  
 235. *Subject Headings*  
 236. *Summary*  
 237. *References*  
 238. *Appendix*  
 239. *Notes*  
 240. *Tables*  
 241. *Figures*  
 242. *Supplementary Materials*  
 243. *Index*  
 244. *Abstract*  
 245. *Keywords*  
 246. *Subject Headings*  
 247. *Summary*  
 248. *References*  
 249. *Appendix*  
 250. *Notes*  
 251. *Tables*  
 252. *Figures*  
 253. *Supplementary Materials*  
 254. *Index*

**TAYLOR & FRANCIS**

1

2/27/81

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS  
 SOFTWARE: Wordperfect  
 AT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/011,436  
 FILING DATE: 11-Dec-2001  
 CLASSIFICATION: 435  
 APPLICATION DATA:  
 APPLICATION NUMBER: 08/511,011  
 FILING DATE: 5-AUGUST-1995  
 APPLICATION NUMBER: 08/054,714  
 FILING DATE: 28-APRIL-1993  
 APPLICATION NUMBER: 07/994,928  
 FILING DATE: 22-DECEMBER-1992

KEY/AGENT INFORMATION:  
 NAME: Mary Anne Schofield

REGISTRATION NUMBER: 36,669  
 REFERENCE/DOCKET NUMBER: LUD 5299.9 DIV CON

COMMUNICATION INFORMATION:  
 TELEPHONE: (202) 662-0200  
 TELEFAX: (202) 662-4643

FOR SEQ ID NO: 1:

ICE CHARACTERISTICS:

LENGTH: 1906 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

ACE DESCRIPTION: SEQ ID NO: 1:

1

Similarity 39.0%; Score 30; DB 4; Length 1906;

Conservative 67.9%; Pred. No. 0.032;

Mismatches 24; Indels 1; Gaps 1;

ACTTACNAGCCAGCCAGCATTCTTCTGCGATCCTCTGCGATG-GTCAGGTATTGGAGG 59

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

ACTTACTAGCCAGCCAGCATTCTTCTGCGATGCTCTGCGATGCTGTCAGCGATTGGAGG 865

AGNACACAGCCAGCAGT 77

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

GTACACAGCCATCAGT 883

Application US/08370909

43648

INATION:

ROBBINS, PAUL F.; ROSENBERG,

STEVEN A.

INVENTION: p15 AND TYROSINASE

INVENTION: MELANOMA ANTIGENS AND THEIR USE IN DIAGNOSTIC

SEQUENCES: 23

ENCE ADDRESS:

345 PARK AVENUE

NEW YORK

NEW YORK

USA

154

READABLE FORM:

TYPE: FLOPPY DISK

IBM PC COMPATIBLE

RG SYSTEM: PC-DOS/MS-DOS

ASCII

PLICATION DATA:

ION NUMBER: US/08/370,909

DATE: 10-JAN-1995

CATION: 514

AGENT INFORMATION:

AROL M. GRUPPI

ATION NUMBER: 37,341

RE/DOCKET NUMBER: 2026-4155

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

TELEX: 421792

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 1910

TYPE: NUCLEOTIDE

STRANDEDNESS: DOUBLE

TOPOLOGY: UNKNOWN

MOLECULE TYPE: PEPTIDE

US-08-370-909-18

Query Match 39.0%; Score 30; DB 2; Length 1910;

Best Local Similarity 67.9%; Pred. No. 0.032;

Mismatches 24; Indels 1

Matches 53; Conservative 0;

1 ACTTACNAGCCAGCCAGCATTCTTCTGCGATCCTCTGCGATG-GTCAGGTCA

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

810 ACTTACTAGCCAGCCAGCATTCTTCTGCGATGCTCTGCGATGCTGTCAGCGG

60 AGNACACAGCCAGCAGT 77

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

870 AGTACACAGCCATCAGT 887

RESULT 27

US-08-181-471-1

Sequence 1, Application US/08181471

Patent No. 5641508

GENERAL INFORMATION:

APPLICANT: Li, Lingna

APPLICANT: Lishko, Valeryi K.

TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL

TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Thomas Fitting

STREET: 12526 High Bluff Drive, Suite 300

CITY: San Diego

STATE: CA

COUNTRY: USA

ZIP: 92130

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/181,471

FILING DATE: 13-JAN-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/041,553

FILING DATE: 02-APR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas

REGISTRATION NUMBER: 34,163

REFERENCE/DOCKET NUMBER: ANT0029P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-792-3680

TELEFAX: 619-792-8477

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2384 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: CDS





Query Match 36.9%; Score 28.4; DB 1; Length 3365;

milarity 96.7%; Pred. No. 0.17;  
 Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 TAGCATCATCCTCTGCATGGTCAGGTCAT 52  
 |||||  
 GAGCATCATCCTCTGCATGGTCAGGTCAT 1394  
 1  
 plication US/09446402A  
 23003  
 MATION:  
 lack Jr., Charles A.  
 ENTION: COMPOSITIONS AND METHODS FOR ACTIVATING  
 ENTION: GENES OF INTEREST  
 CE: 5722-2(35722/191928)  
 ICATION NUMBER: US/09/446,402A  
 NG DATE: 1999-12-20  
 ATION NUMBER: PCT/US98/13093  
 DATE: 1998-06-24  
 ATION NUMBER: 60/050,772  
 DATE: 1997-06-25  
 Q ID NOS: 19  
 stSEQ for Windows Version 4.0  
 9  
 rtificial Sequence  
 MATION: Recombinant Molecule containing multiple cloning  
 MATION: site, kozak sequence, lacZ gene.  
 isc\_feature  
 1)...(64)  
 MATION: Multiple cloning site  
 isc feature  
 65)...(79)  
 MATION: Consensus sequence for the "Kozak sequence"  
 MATION: (translation initiation)  
 rim\_transcript  
 80)...(4279)  
 MATION: Beta galactosidase  
 1  
 milarity 36.9%; Score 28.4; DB 4; Length 4279;  
 milarity 96.7%; Pred. No. 0.19;  
 Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 TAGCATCATCCTCTGCATGGTCAGGTCAT 52  
 |||||  
 GAGCATCATCCTCTGCATGGTCAGGTCAT 1864  
 1  
 Application US/08852629  
 06825  
 RMATION:  
 Moyer, Richard W  
 Li, Yi  
 Hall, Richard L  
 NVENTION: ENTOMPOXVIRUS-VERTEBRATE GENE DELIVERY  
 NVENTION: VECTOR AND METHOD  
 SEQUENCES: 17  
 ENCE ADDRESS:  
 E: Saliwanchik, Lloyd & Saliwanchik  
 2421 N.W. 41st Street, Suite A-1  
 ainesville  
 Florida  
 U.S.A.  
 506  
 ADABLE FORM:  
 VPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 FILING DATE:  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bencen, Gerard H  
 REGISTRATION NUMBER: 35,746  
 REFERENCE/DOCKET NUMBER: UF-184  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 407-426-7500  
 TELEFAX: 407-839-8589  
 INFORMATION FOR SEQ ID NO: 15:  
 LENGTH: 4838 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 Query Match 36.9%; Score 28.4; DB 3; Length 4810;  
 Best Local Similarity 96.7%; Pred. No. 0.19;  
 Matches 29; Conservative 0; Mismatches 1; Indels 0;  
 QY 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52  
 |||||  
 Db 2086 CGAGCATCATCCTCTGCATGGTCAGGTCAT 2115  
 RESULT 36  
 US-08-852-629-15  
 ; Sequence 15, Application US/08852629  
 ; Patent No. 6106825  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Moyer, Richard W  
 ; APPLICANT: Li, Yi  
 ; APPLICANT: Hall, Richard L  
 ; TITLE OF INVENTION: ENTOMPOXVIRUS-VERTEBRATE GENE DELIVERY  
 ; TITLE OF INVENTION: VECTOR AND METHOD  
 ; NUMBER OF SEQUENCES: 17  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
 ; STREET: 2421 N.W. 41st Street, Suite A-1  
 ; CITY: Gainesville  
 ; STATE: Florida  
 ; COUNTRY: U.S.A.  
 ; ZIP: 32606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/852,629  
 ; FILING DATE:  
 ; CLASSIFICATION: 514  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bencen, Gerard H  
 ; REGISTRATION NUMBER: 35,746  
 ; REFERENCE/DOCKET NUMBER: UF-184  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 407-426-7500  
 ; TELEFAX: 407-839-8589  
 ; INFORMATION FOR SEQ ID NO: 15:  
 ; LENGTH: 4838 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: both

both  
TYPE: CDNA  
AL: NO  
P: NO  
S: NO

36.9%; Score 28.4; DB 3; Length 4838;  
milarity 96.7%; Pred. No. 0.19; 1; Indels 0; Gaps 0;

Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGTCAGGTGCAT 52

|||||  
GAGCATCATCCTCTGCATGTCAGGTGCAT 2143

Application US/09654449

64061

MATION:

almer, Michelle A.J.

Gee, Melissa

Tillotson, Bonnie

Chang, Xiac-Jia

ENTION: Receptor Function Assay for G-Protein Coupled Receptors and Orpha

ENTION: Receptors by Reporter Enzyme Mutant Complementation

ICE: 4085-226-27

ICATION NUMBER: US/09/654,449

NG DATE: 2000-09-01

ATION NUMBER: US 60/180,669

DATE: 2000-02-07

Q ID NOS: 5

stSEQ for Windows Version 4.0

0

rtificial Sequence

MATION: Nucleotide sequence for pICAST ALC

DS

1457)...(4486)

36.9%; Score 28.4; DB 4; Length 6700;

milarity 96.7%; Pred. No. 0.22; 1; Indels 0; Gaps 0;

Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGTCAGGTGCAT 52

|||||  
GAGCATCATCCTCTGCATGTCAGGTGCAT 2466

7 Application US/09238356

12683

MATION:

Kingsman, et al

ENTION: Retroviral Vectors

CE: 674523-2006

ICATION NUMBER: US/09/238,356

NG DATE: 1999-01-27

ATION NUMBER: PCT/GB/03876

DATE: 1998-12-22

Q ID NOS: 64

PatentIn version 3.0

2

rtificial Sequence, plasmid

isc feature

1)...(7252)

OTHER INFORMATION: plasmid vector  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: AX003206  
DATABASE ENTRY DATE: 2000-08-24  
RELEVANT RESIDUES: (1)...(7252)  
US-09-238-356-27

Query Match 36.9%; Score 28.4; DB 4; Length 7252;  
Best Local Similarity 96.7%; Pred. No. 0.23;  
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGTCAGGTGCAT 52

|||||  
Db 2076 CGAGCATCATCCTCTGCATGTCAGGTGCAT 2105

RESULT 39

5169760-3

Patent No. 5169760

APPLICANT: WILCON, EDWARD R.

TITLE OF INVENTION: METHOD, VECTORS AND HOST CELLS FOR

THE CONTROL OF EXPRESSION OF HETEROLOGOUS GENES FROM LAC

OPERATED PROMOTERS

NUMBER OF SEQUENCES: 3

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/386,821

FILING DATE: 27-JUL-1989

SEQ ID NO:3

LENGTH: 7366

5169760-3

Query Match 36.9%; Score 28.4; DB 6; Length 7366;  
Best Local Similarity 96.7%; Pred. No. 0.23;  
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGTCAGGTGCAT 52

|||||  
Db 2246 CGAGCATCATCCTCTGCATGTCAGGTGCAT 2275

RESULT 40

US-09-238-356-28/c

Sequence 28, Application US/09238356

Patent No. 6312683

GENERAL INFORMATION:

APPLICANT: Kingsman, et al

TITLE OF INVENTION: Retroviral Vectors

FILE REFERENCE: 674523-2006

CURRENT APPLICATION NUMBER: US/09/238,356

CURRENT FILING DATE: 1999-01-27

PRIOR APPLICATION NUMBER: PCT/GB/03876

PRIOR FILING DATE: 1998-12-22

NUMBER OF SEQ ID NOS: 64

SOFTWARE: PatentIn version 3.0

SEQ ID NO 28

LENGTH: 7387

TYPE: DNA

ORGANISM: Artificial Sequence, primer

FEATURE:

NAME/KEY: misc feature

LOCATION: (1)...(7387)

OTHER INFORMATION: plasmid vector

PUBLICATION INFORMATION:

DATABASE ACCESSION NUMBER: AX003207

DATABASE ENTRY DATE: 2000-08-24

RELEVANT RESIDUES: (1)...(7387)

US-09-238-356-28

Query Match 36.9%; Score 28.4; DB 4; Length 7387;  
Best Local Similarity 96.7%; Pred. No. 0.23;  
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGTCAGGTGCAT 52

.4 08:44:46 2004

us-10-090-326-25.rni

|||||  
GAGCATCATCCTCTGCAIGGTCAGGTGAT 2571

sd: April 13, 2004, 20:16:05  
secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

leic search, using sw model

April 13, 2004, 17:09:09 ; Search time 217.5 Seconds  
(without alignments)  
1503.960 Million cell updates/sec

US-10-090-326-25

1 acttaacagccagcanca.....ggagnacaacagccancagct 77

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

3373863 seqs, 2124099041 residues

hits satisfying chosen parameters: 6747726

.length: 0

.length: 2000000000

Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

N Geneseq 25Jan04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002s:\*
- 7: Geneseqn2003as:\*
- 8: Geneseqn2003bs:\*
- 9: Geneseqn2003cs:\*
- 10: Geneseqn2004s:\*

is the number of results predicted by chance to have a  
ter than or equal to the score of the result being printed,  
ived by analysis of the total score distribution.

#### SUMMARIES

Query	Match	Length	DB	ID	Description
94.8	77	6	ABQ82559	Abq82559	Tyrosinas
39.0	1059	2	AAT29076	Aat29076	Tyrosinas
39.0	1059	2	AAT29077	Aat29077	Tyrosinas
39.0	1059	8	ADB16200	Adb16200	Cleavase
39.0	1059	8	ADB16199	Adb16199	Cleavase
39.0	1587	2	AAT29078	Aat29078	Tyrosinas
39.0	1587	2	AAT29079	Aat29079	Tyrosinas
39.0	1587	4	AAF26897	Aaf26897	Human tyr
39.0	1587	8	ADB16201	Adb16201	Cleavase
39.0	1587	8	ADB16201	Adb16201	Cleavase
39.0	1590	5	AAF88041	Aaf88041	Human tyr
39.0	1886	1	AAN80637	Aan80637	Encodes h
39.0	1894	2	AAQ06552	Aaq06552	Clone BBT
39.0	1894	2	AAT01009	Aat01009	Human tyr
39.0	1894	2	AAT75678	Aat75678	Human SK2
39.0	1906	2	AAQ66548	Aaq66548	Human tyr
39.0	1910	2	AAT35901	Aat35901	Tyrosinas
39.0	1964	7	ABQ83843	Abq83843	Human tyr
39.0	1964	9	ADC09569	Adc09569	Tyrosinas
39.0	2041	4	AAQ60958	Aaq60958	Human can
39.0	2384	2	AAQ72871	Aaq72871	Human tyr
39.0	2384	2	AAQ33316	Aaq33316	Tyrosinas
39.0	2384	2	ABQ76197	Abq76197	Human tum

24	30	39.0	6408	7	ABX10643
25	29.2	37.9	8892	2	AAZ11453
26	28.4	36.9	67	2	AAZ87446
27	28.4	36.9	67	2	AAZ87448
28	28.4	36.9	100	9	ADC060826
29	28.4	36.9	1542	6	ABA99646
30	28.4	36.9	3144	6	ABL61145
31	28.4	36.9	3150	4	AAH45081
32	28.4	36.9	3157	6	ABS70119
33	28.4	36.9	3157	7	ACC00558
34	28.4	36.9	3157	8	ACC71946
35	28.4	36.9	3279	6	AAD22647
36	28.4	36.9	3284	9	ADB89192
37	28.4	36.9	3284	9	ADC24539
38	28.4	36.9	3284	9	ADD67809
39	28.4	36.9	3284	9	ADE10522
40	28.4	36.9	3284	9	ADE11444
41	28.4	36.9	3284	9	ADE12622
42	28.4	36.9	3284	9	ADE12385
43	28.4	36.9	3365	2	AAQ41287
44	28.4	36.9	3396	8	ACF57390
45	28.4	36.9	3652	5	AAD04924

#### ALIGNMENTS

RESULT 1  
ABQ82559  
ID ABQ82559 standard; DNA; 77 BP.  
XX AC ABQ82559;  
XX XX  
DT 18-DEC-2002 (first entry)  
XX XX  
DE Tyrosinase mimic internal control probe SEQ ID NO:25.  
XX XX  
KW Human; CEA; carcinoembryonic antigen; adenocarcinoma; oesophagi  
maligancy; probe; ss.  
XX OS Synthetic.  
XX XX  
PN WO200270751-A1.  
XX PD 12-SEP-2002.  
XX PF 04-MAR-2002; 2002WO-US006504.  
XX PR 02-MAR-2001; 2001US-0273277P.  
XX PA (UYFI-) UNIV PITTSBURGH.  
XX PI Godfrey TE, Luketich JD, Raja S, Kelly LA, Finkelstein SD;  
XX DR WPI; 2002-732795/79.  
XX PT Multiplex PCR method for detecting malignancies, e.g. adenocarc  
the esophagus comprises conducting a PCR amplification on a DN  
a PCR reaction mixture.  
XX PS Claim 47; Page 80; 141pp; English.  
XX CC The present invention describes a multiplex polymerase chain re  
(PCR) (M1) comprising conducting PCR on a DNA sample in a react  
mixture conducted in first and second amplification stages, eac  
or more PCR cycles comprising denaturing, annealing and elongat  
where the elongating step may be conducted at the same temperat  
annealing step. The second amplification stage of (M1) is condu  
different reaction conditions from that of the first amplificat  
to modulate the relative rate of production of the first amplif  
first primer set and a second amplicon by a second primer set c  
first and second amplification stages. Also described: (1) an  
oligonucleotide comprising 15-28 bp or its derivative; (2) intr

stic; (3) rapid detection of a malignancy or of metastasised  
 onoma of the oesophagus; or (4) a cartridge for use in an  
 PCR system. (M1) is useful for detecting malignancies, e.g.  
 the time it takes to carry out a PCR reaction. The present  
 represents an internal control probe, which is used in an  
 om the present invention

17 BP; 20 A; 22 C; 15 G; 16 T; 0 U; 4 Other;  
 94.8%; Score 73; DB 6; Length 77;  
 milarity 100.0%; Pred.No.3.8e-18;  
 Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CTTACNAGCCGAGCANCATCTTAGCATCATCTCTCCATGTCAGTCAATTTGGAGGA 60  
 |||||  
 CTTACNAGCCGAGCANCATCTTAGCATCATCTCTCCATGTCAGTCAATTTGGAGGA 60  
 |||||  
 NACAACAGCCANCAGT 77  
 |||||  
 NACAACAGCCANCAGT 77

standard; DNA; 1059 BP.

6 (first entry)

gene fragment.

t; mutation; cleavage; nuclease; cleavage; Thermus;  
 a; Saccharomyces; Campylobacter; Mycobacterium; Shigella;  
 ccus; identification; detection; ss.

ns.

A1.

6.

5; 95WO-US014673.

4; 94US-00337164.

5; 95US-00402601.

5; 95US-00484956.

5; 95US-00520946.

IRD WAVE TECHNOLOGIES INC.

E. Lyamichev VI, Brow MAD, Oldenburg MC, Heisler LM;  
 live DM;

259862/26.

f nucleic acids to detect mutation(s) - allows detection esp.  
 53 gene, to identify strains of microorganisms and viruses.

; Page 268; 433pp; English.

f nucleic acids using an enzyme, especially a nuclease selected  
 roup consisting of Cleavage (RTM) BN enzyme, Thermus aquaticus  
 rase, Thermus thermophilus DNA polymerase, Escherichia coli  
 the Saccharomyces cerevisiae Rad1/Rad10 complex. The nucleic  
 rate is preferably an oligonucleotide containing a human p53  
 nee or alternatively, microbial gene sequences. Cleavage  
 re compared to the cleavage products of reference gene  
 The method is used for detecting mutation in the human p53  
 identifying strains of microorganisms, especially bacteria  
 rom the group of members of the genera Campylobacter,  
 a, Mycobacterium, Salmonella, Shigella and Staphylococcus. The

CC method may also be used for the identification of viruses, esp  
 CC hepatitis C virus and simian immunodeficiency virus. The human  
 CC gene (both wild type and mutant gene fragments) was used as a  
 CC sequence for the method. This sequence is a fragment of the C;  
 CC gene (Exons 1-4)

SQ Sequence 1059 BP; 298 A; 248 C; 227 G; 286 T; 0 U; 0 Other;

Query Match 39.0%; Score 30; DB 2; Length 1059;  
 Best Local Similarity 67.9%; Pred.No. 0.33;  
 Matches 53; Conservative 0; Mismatches 24; Indels 1.

Qy 1 ACTTACNAGCCGAGCANCATCTTAGCATCATCTCTCCATGTCAGTCAATTTGGAGTCA  
 |||||  
 475 ACTTACTCAGCCGAGCATCATCTCTCTCTTGGCAGATTGCTGTAGCGG

Qy 60 AGNACACAGCCANCAGT 77

Db 535 AGTACAACAGCCATCAGT 552.

RESULT 3

AAT29077

ID AAT29077 standard; DNA; 1059 BP.

XX AC AAT29077;

DT 02-DEC-1996 (first entry)

DE Tyrosinase gene fragment (422 mutant).

XX p53; mutant; mutation; cleavage; nuclease; cleavage; Thermus;  
 XX Escherichia; Saccharomyces; Campylobacter; Mycobacterium; Shig  
 XX Staphylococcus; identification; detection; ss.

OS Homo sapiens.

XX WO9615267-A1.

XX 23-MAY-1996.

XX 09-NOV-1995; 95WO-US014673.

XX 09-NOV-1994; 94US-00337164.

XX 09-MAR-1995; 95US-00402601.

XX 07-JUN-1995; 95US-00484956.

XX 30-AUG-1995; 95US-00520946.

XX (THIR-) THIRD WAVE TECHNOLOGIES INC.

XX Dahlberg JE, Lyamichev VI, Brow MAD, Oldenburg MC, Heisler  
 XX Fors L, Olive DM;

XX WPI; 1996-259862/26.

XX Cleavage of nucleic acids to detect mutation(s) - allows detec  
 XX in human p53 Gene, to identify strains of microorganisms and v  
 XX Example 10; Page 268-269; 433pp; English.

XX Cleavage of nucleic acids using an enzyme, especially a nuclea  
 XX from the group consisting of Cleavage (RTM) BN enzyme, Thermus  
 XX DNA polymerase, Thermus thermophilus DNA polymerase, Escherich  
 XX ExoIII and the Saccharomyces cerevisiae Rad1/Rad10 complex. Th  
 XX acid substrate is preferably an oligonucleotide containing a hu  
 XX gene sequence or alternatively, microbial gene sequences. Clea  
 XX products are compared to the cleavage products of reference ge  
 XX sequences. The method is used for detecting mutation in the hu  
 XX gene; for identifying strains of microorganisms, especially ba  
 XX selected from the group of members of the genera Campylobac  
 XX Escherichia, Mycobacterium, Salmonella, Shigella and Staphyloc  
 XX method may also be used for the identification of viruses, esp  
 XX hepatitis C virus and simian immunodeficiency virus. The human

is used for the ide

Exam] 9 10. Page 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120

on relates to a method of detecting and identifying strains of *Y. enterocolitica* by providing a nuclease and a nucleic acid substrate sequences derived from microorganism(s), treating the substrate to form cleavage structure(s) and reacting the nuclease cleavage structures so that cleavage product(s) are produced. This is used for the identification of strains of microorganisms.





nce or alternatively, microbial gene sequences. Cleavage  
 re compared to the cleavage products of reference gene  
 The method is used for detecting mutation in the human p53  
 identifying strains of microorganisms, especially bacteria  
 rom the group of members of the genera Campylobacter,  
 a, Mycobacterium, Salmonella, Shigella and Staphylococcus.  
 also be used for the identification of viruses, especially  
 C virus and simian immunodeficiency virus. The human tyrosinase  
 wild type and mutant gene fragments) was used as a test  
 or the method. This sequence is a fragment of the tyrosinase  
 s 1-5 of the 422 mutant).

587 BP; 410 A; 386 C; 369 G; 422 T; 0 U; 0 Other;

39.0%; Score 30; DB 2; Length 1587;

ilarity 67.9%; Pred. No. 0.37;

Conservative 0; Mismatches 24; Indels 1; Gaps 1;

TTTACNAGCCAGCAGCATCTTGTGCGATG-GTCAGGTCATTGGAGG 59

||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

TTTATCAGCCAGCATCATCTTCTCTTGGCAGATTGCTGTAGCCGATTGGAGG 841

TTACNAGCCAGCATCT 77

||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

TTACNAGCCAGCATCT 859

andard; cDNA; 1587 BP.

l (first entry)

sinase nucleotide sequence SEQ ID NO:3.

sinase; anti-gray hair agent; hair colour; hair follicle; ss.

is.

18-A.

).

; 99JP-00117006.

; 99JP-00117006.

SEIDO CO LTD.

41361/15.

51344.

the effect of an anti-gray hair agent.

Page 5-7; 11pp; Japanese.

invention describes a method for separating the mRNA encoding  
 in a hair follicle for the assumption of hair colour and for  
 in anti-gray hair agent. The method is useful for screening an  
 hair agent and judging the effect of it. The present sequence  
 an tyrosinase which is given in the exemplification of the  
 rention

587 BP; 408 A; 387 C; 370 G; 422 T; 0 U; 0 Other;

39.0%; Score 30; DB 4; Length 1587;

ilarity 67.9%; Pred. No. 0.37;

Conservative 0; Mismatches 24; Indels 1; Gaps 1;

TTTACNAGCCAGCAGCATCTTGTGCGATG-GTCAGGTCATTGGAGG 59

||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 782 ACTTACTCAGCCAGCATCTTCTCTTGGCAGATTGCTGTAGCCGA

QY 60 AGNACAACAGCCAGCATCT 77

||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 842 AGTACACAGCCAGCATCT 859

RESULT 9

ADB16202

ID ADB16202 standard; DNA; 1587 BP.

XX

AC ADB16202;

DT 20-NOV-2003 (first entry)

XX

DE Cleavase BN DNA substrate #12.

XX

KW ds; DNA polymerase; microorganism strain identification; bacte

KW Campylobacter; Escherichia; Mycobacterium; Salmonella; Shigell

KW Staphylococcus; virus; hepatitis C virus; simian immunodeficie

KW Mycobacterium tuberculosis; human.

XX Homo sapiens.

OS

XX US2003054338-A1.

XX

PD 20-MAR-2003.

XX

PF 28-AUG-2001; 2001US-00940925.

XX

PR 07-DEC-1992; 92US-00986330.

PR 04-JUN-1993; 93US-00073384.

PR 06-JUN-1994; 94US-00254359.

PR 09-NOV-1994; 94US-00337164.

PR 09-MAR-1995; 95US-00402601.

PR 07-JUN-1995; 95US-00484956.

PR 30-AUG-1995; 95US-00520946.

PR 06-FEB-1997; 97US-00789079.

PR 19-FEB-1997; 97US-00802233.

PR 05-SEP-2000; 2000US-00655378.

XX

(DAHL/) DAHLBERG J E.

(BROW/) BROW M A D.

(LYAM/) LYAMICHEV V I.

PI Dahlberg JE, Brow MAD, Lyamichev VI;

DR WPI; 2003-615811/58.

XX

PT Identification of strains of microorganisms, by treating nucle:  
 cleavage structure(s) derived from microorganisms with nucleas:  
 cleavage products(s) and detecting the product(s).

XX Example 12; Page 133; 303pp; English.

CC The invention relates to a method of detecting and identifying  
 microorganisms by providing a nuclease and a nucleic acid subst  
 containing sequences derived from microorganism(s), treating th  
 acid substrate to form cleavage structure(s) and reacting the r  
 with the cleavage structures so that cleavage product(s) are pr  
 CC The method is used for the identification of strains of microo  
 CC The microorganism comprises bacteria including Campylobacter,  
 CC Escherichia, Mycobacterium, Salmonella, Shigella or Staphylococ  
 CC virus comprising hepatitis C virus or simian immunodeficiency v  
 CC Mycobacterium comprises strains of multi-drug resistant Mycobac  
 CC tuberculosis. The method is less sensitive to size so that enti  
 CC rather than gene fragments, may be analysed. It facilitates the  
 CC internal standards for subsequent analysis and data comparison,  
 CC increases the productivity of personnel and equipment. The pres  
 CC sequence represents a Cleavase BN substrate DNA.

XX Sequence 1587 BP; 410 A; 386 C; 369 G; 422 T; 0 U; 0 Other;



TTACNCAGCCGAGCAGCAGCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAATTGGAGG 59

b  
848 AGTACAACAGCCATCAGT 865



Melanoma; tyrosinase; immunogen; vaccine; cancer; immunotherapy  
transgenic animal; ds.

Homo sapiens.

[illegible]

```
CDS
29. .1618
/*tag= a
```

WO9621734-A2.

18-JUL-1996.

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

11-JAN-1996; 96WO-US000473.

10-JAN-1995; 95US-00370909.

(USSH ) US DEPT HEALTH & HUMAN SERVICES.

Robbins PF, Rosenberg SA:

WPI; 1996-342287/34.

Nucleic acid sequence encoding p15 melanoma antigen - and immunopeptide(s) derived from it, useful for diagnosis, prevention of melanoma.

A DNA fragment (AA335901) codes for tyrosinase (AAW033306), an enzyme involved in melanin synthesis. The tyrosinase was recognised by a restricted tumour infiltrating lymphocytes from a melanoma patient. This DNA fragment can be used for recombinant production of tyrosinase. Immunogenic peptides (see also AAW03304-05) for use as melanoma antigens are also available.

A DNA fragment (AAT35901) codes for tyrosinase (AAW033306).

involved in melanin synthesis. The tyrosinase was recognised

restricted tumour infiltrating lymphocytes from a melanoma patient. The DNA fragment can be used for recombinant production of tyrosinase (immunogenic potential) and also for immunisation of mice.

immunogenic peptides (see also AAU03304-05) for use as melanoma:

Sequence 1910 BP; 502 A; 451 C; 433 G; 524 T; 0 U; 0 Other;

[illegible]

Only Maccn	39.0%;	Score 30;	DB 2;	Length 1910;
at Local Similarity	67.9%;	Pred No	0.4.	

ches 53; Conservative 0; Mismatches 24; Indels 1;

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

I ACTTACNCAGCCAGCANCACTCTAGCATCATCTCTGCATG-GTCAGGTCAAT-

810 ACTTACTCAGCCGAGCATCATCTTCTCTCTTGGCAGATTGCTGTAGCCGA:



ic acid, used as a marker to determine the effectiveness of , to treat cancer cell growth in individuals.

.ge 460-461; 527pp; English.

on relates to 1046 novel nucleic acids which are used as determining the sensitivity of a cancer cell to the agent TAXOL. Cancer cells can be treated with TAXOL when they express one of the 242 sensitivity markers or the cells are express one of the 804 resistance markers. The methods can determine the effectiveness of TAXOL in the treatment of growth in an individual. The markers can be used as targets using anti-cancer agents such as chemotherapeutic compounds. The agents also be used as targets in developing treatments for cancer, by those cancers which display resistance to agents and exhibit one of the markers. The anticancer agents developed by the novel can be used to treat cancer. Probes based on the markers can be used to detect transcripts or genomic sequences corresponding to the markers. The identification of cells or tissues which mis-express the markers which may be targeted include carcinoma (e.g. squamous cell carcinoma), sarcoma (e.g. fibrosarcoma) leukemia (e.g. lymphocytic leukemia), plasmacytoma, reticulum cell sarcoma, Hodgkin's disease, and other cancers (e.g. glioma). The present sequence is one of the markers for cancer cell markers.

141 BP; 564 A; 467 C; 455 G; 548 T; 0 U; 7 Other;

39.0%; Score 30; DB 4; Length 2041;

Conservative	0; Mismatches	24; Indels	1; Gaps
Conservative	0; Mismatches	24; Indels	1; Gaps

'TTACNCAGCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAITTTGGAGG 59

TTACTCAGCCAGCATCTCTCTCCTTGGCAGATTGCTGTAGCCGATTGGAGG 966

5'NACAACAGCCANCAGT 77

|||||TACAAAGCCATCAGT 984

andard; cDNA: 2384 BP.

```
(revised)
(first entry)
```

inase gene.

sinase; compound; hair; follicle; liposome; lipophilic; stratum corneum; melanin; hair dye; glycoprotein; alopecia; Y; transforming growth factor; growth stimulant; aromatase; A; macromolecule; polymer; ss.

•  
ທ່

```
Location/Qualifiers
503. .2092
/*tag= a
/product= "human tyrosinase"
```

1.

; 94WO-US003634.

; 93US-00041553

; 94US-00181471.

ICANCER INC.

XX	
PI	
XX	Li L, Lishko VK;
XX	
DR	WPI; 1994-322816/41.
DR	P-PSDB; AAR63623.
XX	
PT	Liposomes for delivering protein, nucleic acid etc. to hair fol-
PT	e.g. to restore hair colour, prevent hair loss during chemothe-
PT	stimulate hair growth etc.
PT	

Claim 6; Page 67-70; 100pp; English.

The nucleotide sequence of a human tyrosinase gene. This is an a compound which can be delivered to hair follicles via a novel composition. The liposomes are comprised of a lipophilic or lipophobic compound which will selectively target the hair follicle (via a compound) without damaging or unwanted effects on cells outside of the hair follicle. Compounds e.g. tyrosinase (or the DNA encoding it), or hair dyes, can be delivered to the hair follicle to restore hair or condition. Other compounds targeted at hair follicles can include glycoprotein (AAQ72872) (for treatment of chemotherapy-induced human transforming growth factor alpha (AAQ72873) (for reversal hair); hair growth stimulants (e.g cyclosporin A or aromatase) antisense sequences. The methods allows compounds (e.g. macromol polymers), which would not normally reach the hair follicles, to be delivered to these target areas. (Updated on 25-MAR-2003 to correct field.)

Sequence 2384 BP; 652 A; 534 C; 519 G; 679 T; 0 U; 0 Other;

very Match 39.0%; Score 30; DB 2; Length 2384;

Best Local Similarity 67.9%; Pred. No. 0.43;

Matches	53;	Conservative	0;	Mismatches	24;	Indels	1;
---------	-----	--------------	----	------------	-----	--------	----

QY 1 ACTTACNCAGCCAGCANCATTTCTAGCATCATCCTCTGCATG-GTCAGGTCATT

D<sub>b</sub> 1284 ACTTACTCAGCCCGCATCTTCTCTCTTGGCAGATTGTCTGTAGCCGAT

QY 60 AGNACACAGCCANCAGT 77

Db 1344 AGTACAACAGCCATCAGT 1361

RESULT 22

AB001 22  
AAT33316  
ID AAT33316 standard; cDNA to mRNA; 2384 BP.

AC AAT33316;

DT	18-OCT-1996	(first entry)
AA		

Tyrosinase mRNA.

Human; tyrosinase; detection; primer; reverse transcription; an transcription region; detection; ds.

OS Homo sapiens.

XX	Key	Location/Qualifiers
FH	CDS	503. .2092
FT		

```
FI      /*tag= a
PT      /product= "Tyrosinase"
```

```

FT primer_bind
790. .810
FT **ad- b

```

```
FT primer bind /note= "Binds primer given in AAT33323"
FT primer bind 001 001
```

```

ET /*tag= c
ET /tag= c
ET "P"
ET /tag= c
ET

```

FT	primer_bind	1297.	.131
FT	primer_bind	1297.	.131

```
FT /note= "Bi
```

```
ET
/*tag= e
```

1 /note= "Binds primer given in AAT33319"  
complement(1521..1541)  
/\*tag= f  
3 /note= "Binds primer given in AAT33318"  
complement(1600..1620)  
/\*tag= g  
3 /note= "Binds primer given in AAT33322"  
complement(1650..1670)  
/\*tag= h  
3 /note= "Binds primer given in AAT33320"  
complement(1673..1693)  
/\*tag= i  
/note= "Binds primer given in AAT33324"

-A.

5.  
4; 94JP-00288041.

4; 94JP-00288041.

LA CHEM IND INC.

316329/32.  
N00184.

of tyrosinase mRNA - by amplification of tyrosinase mRNA allows  
selection of trace amounts of RNA.

; Page 7-9; 10pp; Japanese.

ace encodes human tyrosinase. This sequence was detected by the  
the invention using the primers given in AAT33317-24. The  
prises synthesizing tyrosinase cDNA from a sample by reverse  
ion and then amplifying the cDNA using primers specific for, or  
ary to, sequences within the transcription region of the  
gene. Amplification using these primers produces a competitive  
can also be amplified. This method can be used to specifically  
ce amounts of human tyrosinase mRNA

384 BP; 652 A; 533 C; 520 G; 679 T; 0 U; 0 Other;

ilarity 39.0%; Score 30; DB 2; Length 2384;

Conservative 0; Mismatches 24; Indels 1; Gaps 1;

TTTACNAGCCAGCCAGCAGCATCTCTGCATG-GTCAGGTCATTGGAGG 59  
|||||  
TTTACTCAGCCAGCAGCATCTCTCTCTTGGCAGATTGCTGTAGCCGATTGGAGG 1343  
|||||

3NACACAGCCAGCAGT 77

|||||  
3TACACAGCCATCAGT 1361

tandard; DNA; 2384 BP.

2 (first entry)

ur antigen tyrosinase DNA.

igen; human; vaccine; cellular immune response; immunogen;  
mour; tyrosinase; ds.

ns.

Bl.

PD 11-SEP-2001.  
XX  
PF 06-APR-1998; 98US-00056105.  
XX  
PR 10-APR-1997; 97US-0043467P.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Kipps TU, Wu Y;  
XX  
DR WPI; 1998-583198/49.  
XX  
PT Generating cellular immune response in patient to target prote  
PT comprises introducing vector with nucleotide sequence encoding  
PT comprising protein processing signal into cell of patient.  
XX  
PS Disclosure; Col 19-22; 61pp; English.  
XX  
CC This invention describes a novel method for generating a cellu  
CC response in a patient to a target protein or its fragment. The  
CC involves introducing a vector containing a nucleotide sequence  
CC chimeric immunogen comprising a protein processing signal and  
CC protein or its fragment. The immunogen is produced by the cell  
CC processed so that the target protein or its fragment is presen  
CC patients immune system and a cellular immune response is initi  
CC method and vectors can be used as a form of vaccination and co  
CC to generate a cellular immune response in patients to, e.g. ca  
CC tumors. The cellular immune response is the predominant immun  
CC in the patient. This sequence represents a DNA fragment which  
CC human tumour antigen tyrosinase described in the method of the  
CC Note: The information in this spec has been previously disclos  
CC WO199845444 however this spec contained no sequence informatio  
XX  
SQ Sequence 2384 BP; 652 A; 533 C; 520 G; 679 T; 0 U; 0 Other;  
Query Match 39.0%; Score 30; DB 2; Length 2384;  
Best Local Similarity 67.9%; Pred. No. 0.43;  
Matches 53; Conservative 0; Mismatches 24; Indels 1;  
QY 1 ACTTACNAGCCAGCCAGCAGCATCTCTGCATG-GTCAGGTCAT  
1284 ACTTACTCAGCCAGCAGCATCTCTCTCTTGGCAGATTGCTGTAGCCGA  
DB 60 AGNACACAGCCAGCAGT 77  
1344 AGTACACAGCCATCAGT 1361  
QY  
DB  
RESULT 24  
ABX10643  
ID ABX10643 standard; DNA; 6408 BP.  
XX  
AC ABX10643;  
XX  
DT 22-APR-2003 (first entry)  
XX  
DE Vector, htyr-pING, containing human tyrosinase gene.  
XX  
KW Human; gene; ds; melanoma; xenogeneic differentiation antigen;  
KW tyrosinase; canine; gene therapy; dog; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN US2002150589-A1.  
XX  
PD 17-OCT-2002.  
XX  
PF 27-NOV-2001; 2001US-00996128.  
XX  
PR 10-DEC-1996; 96US-0032535P.  
PR 18-FEB-1997; 97US-0036419P.  
PR 10-DEC-1997; 97WO-US022669.  
PR 21-MAY-1999; 99US-00308697.



1; 2000US-0180651P.  
1; 2000US-00627694.

IGHTON A N.  
IGMAN P J.  
CHOK J D.

1; Bergman PJ, Wolchok JD;

82484/18.

lanoma in a mammalian subject comprises administering to the immunological amount of a xenogeneic differentiation antigen type as a differentiation antigen expressed by melanoma cells ect.

age 5-8; 15pp; English.

on discloses a method for treating melanoma in a mammalian ch comprises administering to the subject an immunological xenogeneic differentiation antigen (DA) of the same type as a ad by melanoma cells of the subject. Also disclosed are pING itaining a sequence encoding the human or mouse tyrosinase methods and xenogeneic DA are useful for treating canine melanoma (gene therapy) in dog suffering from the disease by ng an immunological amount of the xenogeneic DA and for other ie sequence presented is the vector, htyr-pING, containing the inase gene

08 BP; 1621 A; 1583 C; 1509 G; 1695 T; 0 U; 0 Other;

ilarity 67.9%; Score 30; DB 7; Length 6408;

Conservative 0; Mismatches 24; Indels 1; Gaps 1;

TTACNAGCCAGCCAGCAGCATCTAGCATCATCTCTGCATG-GTCAGGTGATTTGGAGG 59  
|||||  
TTACTCAGCCAGCAGCATCTCTCTCTGGCAGATTGCTGTAGCCGATTTGGAGG 5356

NACAAAGCCAGCAGT 77

TACAAAGCCATCAGT 5374

andard; DNA; 8892 BP.

(first entry)

vector pEUCUT-2 sequence.

vector; functional splice donor site; hybrid viral vector; splice acceptor site; in vivo gene delivery; therapeutic; vector; modified hematopoietic stem cell; MHSC; tumour; MLV; hypoxia response element; HRE; hypoxia; promoter; EIAV; ss.

2.

1.

1; 98WO-GB002885.

1; 97GB-00020216.

1; 97GB-00020465.

ORD BIOMEDICA UK LTD.

inley KM, Bebbington C, Naylor S;

XX  
DR WPI; 1999-263482/22.  
XX  
PT New retroviral vectors, for, e.g. delivering nucleotide sequen  
XX solid tumor sites.  
XX  
PS Example 2 (page 172); Fig 11 (Page 12-13/35); 288pp; English.  
XX  
CC The invention relates to a retroviral vector (RVV) comprising  
CC functional splice donor site (FSDS) and a functional splice ac  
CC (FSAS) where: (i) the FSDS and the FSAS flank a first nucleoti  
CC of interest (NOI); (ii) the FSDS is upstream of the FSAS; (iii)  
CC is derived from a retroviral pro-vector; (iv) the retroviral p  
CC comprises a first nucleotide sequence (NS) capable of yielding  
CC and a second NS capable of yielding the FSAS; and (v) the first  
CC downstream of the second NS, such that the RVV is formed as a  
CC reverse transcription of the retroviral pro-vector. A hybrid v  
CC (VV) system for in vivo gene delivery, which system comprises  
CC VV which encodes a secondary VV, the primary vector capable of  
CC a first target cell and of expressing the secondary VV, which  
CC vector is capable of transducing a secondary target cell, where  
CC primary vector is obtainable from or is based on a adenoviral  
CC the secondary VV is obtainable from or is based on a adenoviral  
CC lentiviral vector (LVV) is also provided. The systems can be u  
CC delivering NOIs to one or more target sites. The NOIs may encod  
CC therapeutic or diagnostic agents. The methods are used particu  
CC producing modified hematopoietic stem cells (MHSCs) to deliver  
CC sites such as solid tumours which are characterised by ischemi  
CC hypoxia or low glucose concentration. The system permits the st  
CC expression of NOIs in targeted cells, e.g. rapidly dividing cel  
CC present sequence represents a EIAV lentiviral version of the v  
CC (Intron Created Upon Transduction). The present plasmid is pEUC  
XX  
SQ Sequence 8892 BP; 2116 A; 2233 C; 2345 G; 2198 T; 0 U; 0 Other;

Query Match 37.9%; Score 29.2; DB 2; Length 8892;  
Best Local Similarity 91.2%; Pred. No. 1.4;  
Matches 31; Conservative 0; Mismatches 3; Indels 0;

OY 19 CATTCTAGCATCATCTCTGCATGTGTCAGGTGAT 52  
DB 2540 CGTAGGAGCATCATCTCTGCATGTGTCAGGTGAT 2573

RESULT 26

AAx87446

ID AAX87446 standard; DNA; 67 BP.

AC AAX87446;

DT 08-OCT-1999 (first entry)

DE Plasmid pJP5603 PCR primer b-left.

KW ET cloning, recE; recT; Escherichia coli; PCR; primer; pJP5603;  
KW homologous recombination; ss.

OS Synthetic.

PN WO9929837-A2.

PD 17-JUN-1999.

PF 07-DEC-1998; 98WO-EP007945.

PR 05-DEC-1997; 97EP-00121462.

PR 05-OCT-1998; 98EP-00118756.

PA (EMBL-) EMBL EURO LAB MOLEKULARBIOLOGIE.

PI Stewart F, Zhang Y, Buchholz F;

XX WPI; 1999-457893/38.

DR

or cloning DNA using a recE and recT homologous recombinant

Page 137; 139pp; English.

a nucleotide sequence of a primer used in the PCR amplification  
; DNA fragment used in a method of the invention. The template  
; and the targeting vector was pSV-paz11. Host cells and/or  
pressing Escherichia coli recE and recT genes (or functionally  
ies) are useful in a novel cloning method, designated ET  
ich involves homologous recombination via the recET mechanism  
achromosomal DNA fragment (the fragment to be cloned), with a  
erably circular DNA vector. The method is particularly  
to complex or difficult DNA constructs, such as those intended  
ous recombination in eukaryotic cells

7 BP; 19 A; 15 C; 21 G; 12 T; 0 U; 0 Other;

ilarity 36.9%; Score 28.4; DB 2; Length 67;

Conservative 96.7%; Pred. No. 0.53; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCTCTGCATGTCAGTGCAT 52

TAGCATCATCTCTGCATGTCAGTGCAT 32

standard; DNA; 67 BP.

(first entry)

95603 PCR primer c-left.

recE; recT; Escherichia coli; PCR; primer; pJP5603;  
recombination; ss.

42.

3.

3; 98WO-EP007945.

7; 97EP-00121462.

3; 98EP-00118756.

3L EURO LAB MOLEKULARBIOLOGIE.

Zhang Y, Buchholz F;

457893/38.

or cloning DNA using a recE and recT homologous recombinant

Page 137; 139pp; English.

e nucleotide sequence of a primer used in the PCR amplification  
r DNA fragment used in a method of the invention. The template  
3 and the targeting vector was pSV-paz11. Host cells and/or  
pressing Escherichia coli recE and recT genes (or functionally  
nes) are useful in a novel cloning method, designated ET  
hich involves homologous recombination via the recET mechanism  
AX87427-29). The method covalently combines one preferably  
trachromosomal DNA fragment (the fragment to be cloned), with a  
ferably circular DNA vector. The method is particularly

CC applicable to complex or difficult DNA constructs, such as tho;  
CC for homologous recombination in eukaryotic cells

XX Sequence 67 BP; 19 A; 16 C; 18 G; 14 T; 0 U; 0 Other;  
SQ

Query Match 36.9%; Score 28.4; DB 2; Length 67;

Best Local Similarity 96.7%; Pred. No. 0.53; Mismatches 1; Indels 0;

Matches 29; Conservative 0; Mismatches 1; Indels 0;

QY 23 CTAGCATCATCTCTGCATGTCAGTGCAT 52

Db 11 CGAGCATCATCTCTGCATGTCAGTGCAT 40

RESULT 28

ADC60826

ID ADC60826 standard; DNA; 100 BP.

XX AC ADC60826;

XX 18-DEC-2003 (first entry)

DE Beta-galactosidase reporter sequence #11.

XX Activated transcription factor; cis element; reporter sequence.

KW beta-galactosidase; gene; ds.

XX OS Escherichia coli.

PN US2003143547-A1.

XX 31-JUL-2003.

XX 24-JAN-2002; 2002US-00057810.

XX 24-JAN-2002; 2002US-00057810.

PA (LIXX/) LI X.

PA (JIAN/) JIANG X.

PI Li X, Jiang X;

DR WPI; 2003-744884/70.

Identifying multiple different activated transcription factors  
sample by transducing or transfecting a cell sample and determ  
activated transcription factors are present in the cell sample

Disclosure; SEQ ID NO 41; 34pp; English.

The invention relates to a method for identifying multiple dif  
activated transcription factors in a cell sample, comprising t  
or transfecting a cell sample to comprise a library of constru  
forming mRNA transcription products by those of the transduced  
transfected cells in which an activated transcription factor i  
that binds to the cis element of the construct present in the  
activates transcription of the reporter sequence of the constr  
in the cell, determining which reporter sequences are compris  
the mRNA transcription products and determining which activate  
transcription factors are present in the cell sample, based on  
reporter sequences were transcribed. Each construct comprises  
element sequence comprising one or more copies of a cis element  
a transcription factor is capable of binding, the cis element  
varying within the library of constructs, a promoter sequence  
to the cis element sequence and a reporter sequence 3' relativ  
promoter sequence that varies within the library, where the sa  
element sequence is employed with a given reporter sequence wi  
library of constructs. The method is useful for identifying mul  
different activated transcription factors in a cell sample for  
a disease state. This sequence represents a reporter sequence  
method of the invention.

Sequence 100 BP; 22 A; 24 C; 32 G; 22 T; 0 U; 0 Other;

36.9%; Score 28.4; DB 9; Length 100;  
ilarity 96.7%; Pred. No. 0.61;  
Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
TAGCATCATCTCTGTCATGGTCAGGTCAT 52  
|||||  
TAGCATCATCTCTGTCATGGTCAGGTCAT 76  
|||||  
andard; DNA; 1542 BP.  
; (first entry)  
associated DNA fragment #2.  
immunodominant; glycoprotein D; antigen; serological testing;  
lex virus.  
2.  
.  
; 2000RU-00101648.  
; 2000RU-00101648.  
TOR VIROLOGY & BIOTECHN RES CENTRE.  
MA, Suslopatorov IM, Plyasunov IV;  
94324/25.  
49704.  
plasmid DNA phavdl determining expression of gene US6  
herpes simplex type-1 virus encoding immuno-dominant group of  
n d(gd)hsv-1 in bacterium Escherichia coli cells.  
Col 13-16; 9pp; Russian.  
ion describes a novel recombinant plasmid DNA constructed in  
containing the herpes simplex type-1 virus (HSV-1) US6 gene  
ich encodes an immunodominant group of glycoprotein D (gD).  
pride shows the antigenic properties of herpes simplex virus  
ified recombinant protein can be used as HSV-1 antigen for the  
testing of HSV-1 in clinical practice. This sequence encodes  
mplex type-1 virus US6-associated protein useful to the  
42 BP; 325 A; 461 C; 431 G; 325 T; 0 U; 0 Other;  
ilarity 36.9%; Score 28.4; DB 6; Length 1542;  
Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
AGCATCATCTCTGTCATGGTCAGGTCAT 52  
|||||  
AGCATCATCTCTGTCATGGTCAGGTCAT 1103  
andard; DNA; 3144 BP.  
(first entry)

XX DE beta-galactosidase reporter DNA fragment.  
XX KW Alpha-SMA; smooth muscle actin; screening; smooth muscle cell;  
XX KW myofibroblast; gene therapy; beta-galactosidase; reporter; ds.  
XX OS Unidentified.  
XX PN EP1172375-A1.  
XX PD 16-JAN-2002.  
XX PF 22-DEC-2000; 2000EP-00128446.  
XX PR 11-JUL-2000; 2000DE-01033633.  
XX PR 31-OCT-2000; 2000DE-01053879.  
XX PA (ODEN/) ODENTHAL M.  
XX PI Odenthal M, Jung D;  
XX DR WPI; 2002-149590/20.  
XX PT New nucleic acid containing regulatory region of the smooth mus  
XX PT gene, useful e.g. for manipulating gene expression in smooth m  
XX PS Disclosure; Page 21-22; 44pp; German.  
XX CC This invention describes a novel nucleic acid (I) comprising:  
CC least one functional region (Ia) from the regulatory region of  
CC smooth muscle actin (SMA) gene and (ii) at least one additional  
CC functional sequence (Ib) operably linked to (Ia). The products  
CC invention can be used for preparing genetically modified eukary  
CC or organisms, for isolation and screening of smooth muscle cell  
CC myofibroblasts or related cells, and for manipulation of gene e  
CC and/or cell function in smooth muscle cell or myofibroblasts,  
CC particularly for gene therapy. Component (Ia) provides cell-ty  
CC differentiation-specific expression or modulation of genes. Thi  
CC represents a beta-galactosidase DNA fragment which can be used  
CC reporter molecule under the control of the alpha-SMA described  
CC invention  
XX SQ Sequence 3144 BP; 694 A; 857 C; 908 G; 685 T; 0 U; 0 Other;  
Query Match 36.9%; Score 28.4; DB 6; Length 3144;  
Best Local Similarity 96.7%; Pred. No. 1.9;  
Matches 29; Conservative 0; Mismatches 1; Indels 0;  
Qy 23 CTAGCATCATCTCTGTCATGGTCAGGTCAT 52  
Db 1143 CGAGCATCATCTCTGTCATGGTCAGGTCAT 1172  
|||||  
RESULT 31  
AAH45081  
ID AAH45081 standard; DNA; 3150 BP.  
AC AAH45081;  
XX 04-SEP-2001 (first entry)  
DT LacZ gene.  
XX LacZ; immunosuppressive; epigenetic regulation motif; immune re  
XX T-cell response; methylation activity;  
XX methyl DNA binding protein identification; ds.  
XX Escherichia coli.  
XX PN W0200140478-A2.  
XX PD 07-JUN-2001.

0; 2000WO-EP012793.

9; 99CA-02291367.

IST PASTEUR.  
NT NAT RECH SCI.

Henry I, Chouluka A;

367812/38.

ed polynucleotide having reduced or increased content of  
control motifs for studying, increasing and/or reducing gene  
expression, and improving DNA vaccination methods.

Fig 1; 75pp; English.

The invention relates to modified LacZ genes (see AAH45079 and  
in which epigenetic regulation motifs have been mutated  
of the wild-type sequence (the present sequence). The mutant  
are useful for inducing in a second host, a protective immune  
against a gene product of a first host. The mutant Lac Z genes  
are useful for evaluating a promoter in biological systems, for  
methylation activity in biological systems and/or for  
unknown methyl DNA binding proteins. The mutant LacZ genes  
are useful for compensating a genetic defect, and for therapeutic  
uses. The mutant LacZ genes are also useful for minimising a T-  
cell response against the T-cells or tissues treated with them

150 BP; 706 A; 853 C; 904 G; 687 T; 0 U; 0 Other;

milarity 36.9%; Score 28.4; DB 4; Length 3150;

Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52

|||||  
GAGCATCATCCTCTGCATGGTCAGGTCAT 1178

standard; DNA; 3157 BP.

2 (first entry)

tosidase gene.

tion modulation system; gene expression cassette; promoter;  
action domain; DNA-binding domain; response element; LBD; DBD;  
clear receptor ligand binding domain; gene therapy; TD; ECR;

a coli.

2-A2.

2.

2; 2002WO-US005090.

1; 2001US-0269799P.

1; 2001US-0313925P.

HM & HAAS CO.

Kumar MB, Cress DE, Fujimoto TT;

682765/73.

PT Novel gene expression modulation system useful for modulating  
PT of a gene of interest, comprising a Group H nuclear receptor J  
XX binding domain comprising a substitution mutation.

PS Example 1; Page 118-120; 148pp; English.

XX The present invention relates to a new gene expression modulat  
CC The invention comprises a gene expression cassette, capable of  
CC expressed in a host cell, comprising a polynucleotide that enc  
CC polypeptide comprising a transactivation domain, a DNA-binding  
CC that recognises a response element associated with a gene whose  
CC expression is to be modulated, and a Group H nuclear receptor  
CC binding domain (LBD). The invention is useful for modulating t  
CC expression of the gene in a host cell which involves introduc  
CC cell, and introducing into the host cell a ligand, where  
CC be modulated is a component of a gene expression cassette comp  
CC response element by the DBD (DNA-binding domain), a promoter t  
CC activated by the TD (transactivation domain) and a gene whose  
CC is to be modulated, where upon introduction of the ligand into  
CC cell, expression of the desired gene is modulated. The induci  
CC expression system and its use for modulating gene expression i  
CC cell overcome the limitations of currently available inducible  
CC systems and provides an effective means for controlling gene e  
CC The present nucleic acid sequence represents a DNA sequence th  
CC in the methods of the invention

XX SQ Sequence 3157 BP; 697 A; 856 C; 915 G; 689 T; 0 U; 0 Other;

Query Match 36.9%; Score 28.4; DB 6; Length 3157;

Best Local Similarity 96.7%; Pred. No. 1.9;  
Matches 29; Conservative 0; Mismatches 1; Indels 0;

QY 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52

|||||  
Db 1170 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1199

RESULT 33

ACC00558  
ID ACC00558 standard; DNA; 3157 BP.

XX AC ACC00558;

XX DT 23-JUN-2003 (first entry)

XX DE E. coli lacZ reporter gene.

XX KW Ecdysone receptor; Bacr; pesticide; insect; male sterility; w  
XX lacZ; ds.

XX OS Escherichia coli.

XX PN WO2003027266-A2.

XX PD 03-APR-2003.

XX PF 20-FEB-2002; 2002WO-US005234.

XX PR 26-SEP-2001; 2001US-0325534P.

XX PA (ROHM ) ROHM & HAAS CO.

XX PI Zhang J, Cress DE, Palli SR, Dhadialla TS;

XX DR WPI; 2003-342734/32.

XX Novel isolated whitefly (Bemisia argentifolii) ecdysone recepto  
PT polypeptide, useful for screening for molecules e.g., agonist  
PT antagonist of ecdysone receptor activity in a cell, which are  
PT pesticides.

XX Example 2; Page 76-78; 85pp; English.

ion relates to a whitefly (*Bemisia argentifolii* edysone BaEcr)) polypeptide and encoding polynucleotides. The BaEcr and polynucleotides are useful for modulating gene expression of some receptor-based gene expression system, and for identifying ligands specific for binding to a ligand-binding domain of the receptor. The polynucleotides are useful for identifying and selecting compounds exhibiting specific binding to the ligand binding domain of BaEcr. Agonists and antagonists modulate insect physiology and development, thus modifying the normal sequence of developmental agents in an insect. For accelerating insect development can be achieved for instance in sterile males. Alternatively, it may be useful slow development of insects, such that the insects reach destructive stages of pestation, such that commercial crops may have passed sensitive stages. Antagonists and antagonists are also useful for artificially maintaining a specific developmental stage. The present sequence relates to an E. coli lacZ reporter gene, used in the construction of BaEcr gene expression cassettes

157 BP; 697 A; 856 C; 915 G; 689 T; 0 U; 0 Other;

Query Match 36.9%; Score 28.4; DB 7; Length 3157;  
Best Local Similarity 96.7%; Pred. No. 1.9;  
Matches 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGTCAGTGCAGGTGCAT 52  
|||||  
TAGCATCATCCTCTGTCAGTGCAGGTGCAT 1199

Standard; DNA; 3157 BP.

3 (first entry)

cZ reporter gene.

; leafhopper; edysone receptor; homopteran; insecticide; ssion; lacZ; gene; ds.

a coli.

39-A1.

3.

2; 2002WO-US005026.

1; 2001US-0325096P.

IM & HAAS CO.

181913/45.

ated leafhopper edysone receptor polypeptide from homopteran hopper e.g. Nephrotetix cincticeps, for identifying molecules the receptor activity and in insecticide discovery.

Page 71-73; 81pp; English.

ion relates to an isolated leafhopper edysone receptor (NcEcr) from the homopteran green leafhopper (Nephrotetix) and encoding polynucleotide. The NcEcr polypeptide, and vectors comprising the polynucleotide, are useful in gene expression and insecticide discovery. The NcEcr and polynucleotide are useful for regulating gene expression of interest in a host cell and in identifying new that modulate activity of a leafhopper edysone receptor (Ecr). Antibodies are useful as a reagent for determining qualitative

CC or quantitative presence of the receptor and to separate or pi  
CC receptor. A composition comprising NcEcr polypeptide, polynuc  
CC vectors are useful for formulation of biological material for  
CC gene expression modulation system or a ligand-screening assay.  
CC Present sequence represents an E. coli lacZ reporter gene frag  
CC in the construction of leafhopper edysone receptor gene expre  
CC cassettes

XX Sequence 3157 BP; 697 A; 856 C; 915 G; 689 T; 0 U; 0 Other;

Query Match 36.9%; Score 28.4; DB 8; Length 3157;  
Best Local Similarity 96.7%; Pred. No. 1.9;  
Matches 29; Conservative 0; Mismatches 1; Indels 0;

QY 23 CTAGCATCATCCTCTGTCAGTGCAGGTGCAT 52

DB 1170 CGAGCATCATCCTCTGTCAGTGCAGGTGCAT 1199

RESULT 35

AAD22647

ID AAD22647 standard; DNA; 3279 BP.

XX

AC AAD22647;

XX 26-FEB-2002 (first entry)

XX

DE Beta-galactosidase reporter gene.

XX

KW Beta-galactosidase; gene expression; human immunodeficiency vi

KW psi gene; nucleocapsid protein; HIV inhibitor screening; ds.

XX

OS Unidentified.

XX

PN WO200177312-A1.

XX

PD 18-OCT-2001.

XX

PF 18-OCT-2000; 2000WO-KR001173.

XX

PR 08-APR-2000; 2000KR-00018489.

XX

PA (YOOJ/) YOO J C.

XX

PI Nam H, Kim S;

XX

DR WPI; 2002-025900/03.

XX

PT Novel microorganism for screening HIV inhibitors, is cotransfo

PT plasmid vector (PV) containing gene expressing HIV nucleocapsi

PT and FV containing HIV psi gene and beta-galactosidase reporter

XX

PS Claim 7; Page 26-29; 35pp; English.

XX

CC The invention relates to a microorganism cotransformed with a  
CC expressing human immunodeficiency virus (HIV) nucleocapsid pro  
CC plasmid vector containing HIV psi gene and beta-galactosidase  
CC gene. The microorganism is useful for screening HIV packaging  
CC by culturing microorganism, treating microorganism with putati  
CC compounds or compositions of HIV inhibitors, and measuring the  
CC change in beta-galactosidase expression in the culture. Microo  
CC useful for high throughput screening (HTS) of HIV inhibitors.  
CC sequence is beta-galactosidase reporter gene which is used for  
CC constructing pNH1 (SL1234) plasmid

XX Sequence 3279 BP; 730 A; 888 C; 940 G; 721 T; 0 U; 0 Other;

Query Match

Best Local Similarity 36.9%; Score 28.4; DB 6; Length 3279;

Matches 29; Conservative 0; Mismatches 1; Indels 0;

QY 23 CTAGCATCATCCTCTGTCAGTGCAGGTGCAT 52

|||||

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCA 52  
|||  
Db 1254 CGAGCATCATCCTCTGCATGGTCAGGTCA 1283

RESULT 37  
ADC24539

ADC24539;

DT 01-JAN-2004 (first entry)

Human beta 2-adrenergic receptor cDNA.

ss; minicell; episomal expression construct; cancer; asthma; inflammation; rheumatoid arthritis; diabetes; Alzheimer's disease; Parkinson's disease; HIV; bacterial infection; hepatitis; myocardial ischaemia; human; gene.

OS Homo sapiens.

OS Homo sapiens.

AA  
PN  
US2003190749-A1

22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
85

XX

28-MAY-2002; 2002US-00157215.  
PF  
XX

PR 24-MAY-2001; 2001US-0293566P.

PR 25--FEB-2002; 2002US-0359843P.

PA (SITRB /) SITRER M W

PA (SABB/) SABBADINI R. A.

PT New minicell-producing parent cell comprising an expression el  
PT mutation in an endogenous gene, useful for producing achromoc  
PT anucleate cells for diagnostic or therapeutic purposes and fo  
PT discovery.

PS Example 22; SEQ ID NO 186; 242pp; English.

The invention relates to a minicell-producing parent cell. The cell comprises: an expression element that comprises a gene or linked to expression sequences that are inducible and/or repressible; where induction or repression of the gene regulates the copy number of episomal expression construct and/or causes or enhances the production of minicells; and/or a mutation in an endogenous gene, where the mutation regulates the copy number of an episomal expression construct and/or causes or enhances minicell production. Also disclosed are compositions and methods for preparing the minicells (or a soluble and/or extractable protein, or antibodies and/or antibody derivatives that recognize an immunogenic epitope on the native form of a membrane protein, associating a radioactive compound with a cell), a method of determining the amount of a membrane protein from a minicell membrane to a biological sample, a pharmaceutical composition comprising the minicell, a method of producing a pharmaceutical composition, a method of detecting an antigen specifically bound by a binding moiety, a method of in situ labeling of a tissue or organ, methods of determining the rate or amount of release of nucleic acid from a minicell to a cell, a method of determining the three-dimensional structure of a membrane protein, a method of identifying ligand-interacting atoms in a defined three-dimensional structure of a target protein, methods of identifying a nucleic acid sequence of a target protein, and methods of bioremediation. The production of the above protein, and methods of bioremediation. The producing parent cell is useful for producing achromosomal and non-replicating cells for diagnostic and therapeutic applications (e.g. in diagnosing cancer, asthma, allergies, inflammation, hemophilia, AIDS, treating cancer, asthma, allergies, inflammation, hemophilia, AIDS).

PS Example 22; SEQ ID NO 186; 242bp; English.

The invention relates to a minicell-producing parent cell. The cell comprises: an expression element that comprises a gene or linked to expression sequences that are inducible and/or repressible; where induction or repression of the gene regulates the copy number of episomal expression construct and/or causes or enhances the production of minicells; and/or a mutation in an endogenous gene, where the mutation regulates the copy number of an episomal expression construct and/or causes or enhances minicell production. Also disclosed are compositions and methods for preparing the minicells (or a soluble and/or extractable protein, or antibodies and/or antibody derivatives that recognize an immunogenic epitope on the native form of a membrane protein, associating a radioactive compound with a cell), a method of determining the amount of a membrane protein from a minicell membrane to a biological material, a pharmaceutical composition comprising the minicell, a method of producing a pharmaceutical composition, a method of detecting an antigen in a sample by immunoassay, a method of detecting an antigen specifically bound by a binding moiety, a method of in situ labeling of a tissue or organ, methods of determining the rate or amount of release of a nucleic acid from a minicell to a cell, a method of determining the three-dimensional structure of a membrane protein, a method of identifying ligand-interacting atoms in a defined three-dimensional structure of a target protein, methods of identifying a nucleic acid sequence of a target protein, and methods of bioremediation. The invention also provides a minicell-producing parent cell useful for producing achromosomal and non-replicating cells for diagnostic and therapeutic applications (e.g. in diagnosing cancer, asthma, allergies, inflammation, hemophilia, AIDS, treating cancer, asthma, allergies, inflammation, hemophilia, AIDS).

PS Example 22; SEQ ID NO 186; 242bp; English.

Alzheimer's disease, Parkinson's disease, HIV, bacterial hepatitis or myocardial ischaemia), as well as research tools for drug discovery or for delivery of nucleic acids and other compounds to cells. The present sequence is a human cDNA used to construct a mini-cell of the invention. Note: The sequence was mixed up the seq id numbers between the disclosure and the invention. This means that several of the sequences cannot be identified and some of the rest may be mis-identified.

284 BP; 753 A; 865 C; 927 G; 739 T; 0 U; 0 Other;

Similarity 36.9%; Score 28.4; DB 9; Length 3284;

Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCTCTGCATGTCAGGTCAT 52  
|||||  
GAGCATCATCTCTGCATGTCAGGTCAT 1283

tandard; DNA; 3284 BP.

4 (first entry)

-2 adrenergic receptor, SalI-XhoI fragment.

ds; membrane protein; transmembrane domain; anchoring domain; Type III secretion system; achromosomal cell; cell; cancer; asthma; allergy; inflammation; arthritis; diabetes; Alzheimer's disease; Parkinson's disease; ion; bacterial infection; hepatitis; myocardial ischaemia;

ns.

14-A1.

3.

2; 2002US-00157299.

1; 2001US-0295566P.

2; 2002US-0359843P.

BBADINI R A.

RKLEY N.

REER M W.

RA, Berkley N, Surber MW;

844449/78.

11 useful for producing achromosomal and anucleate cells for or treating e.g. cancer, asthma, allergies, inflammation, Alzheimer's disease or HIV, and as research tools and agents discovery.

; SEQ ID NO 186; 244pp; English.

ion relates to a minicell comprising at least one nucleic acid. 11 displays a binding moiety directed to a target compound, binding moiety is selected from a eukaryotic membrane protein, arterial membrane protein, an organelle membrane protein, and protein. The fusion protein comprises a first polypeptide at least one transmembrane domain or at least one membrane domain, and a second polypeptide that is not derived from a 1 protein and is neither a His tag nor a glutathione-S- e polypeptide, where the polypeptide comprises a binding so included is the method of introducing a nucleic acid into a

CC cell, comprising contacting the cell with the minicell cited ;  
CC minicell is selected from a eubacterial minicell, a poroplast,  
CC spheroplast and a protoplast. The nucleic acid comprises an e  
CC construct comprising expression sequences operably linked to a  
CC encoding the proteins mentioned above or encoding a therapeutic  
CC polypeptide. The therapeutic polypeptide is a membrane polypep  
CC soluble polypeptide. The soluble polypeptide comprises a cell  
CC secretion sequence. The expression sequences are inducible and  
CC repressible. These are induced and/or depressed when the bind  
CC displayed by the minicell binds to its target compound. The O  
CC polypeptide having an amino acid sequence that facilitates cel  
CC transfer of a biologically active compound contained within o  
CC by the minicell. The membrane of the minicell comprises a syst  
CC transferring a molecule from the interior of a minicell into t  
CC cytoplasm of the cell. The system is a Type III secretion syst  
CC minicell and method are useful in producing achromosomal and a  
CC cells for diagnostic and therapeutic applications (e.g. in dis  
CC treating cancer, asthma, allergies, inflammation, rheumatoid a  
CC diabetes, Alzheimer's disease, Parkinson's disease, HIV, bacte  
CC infections, hepatitis or myocardial ischaemia), as well as res  
CC and agents for drug discovery or for delivery of nucleic acids  
CC bioactive compounds to cells. The present sequence is a human  
CC sequence from a gene of interest, incorporated into a minicell  
CC of the invention.

XX SQ Sequence 3284 BP; 753 A; 865 C; 927 G; 739 T; 0 U; 0 Other;

Query Match 36.9%; Score 28.4; DB 9; Length 3284;  
Best Local Similarity 96.7%; Pred. No. 2;  
Matches 29; Conservative 0; Mismatches 1; Indels 0;

QY 23 CTAGCATCATCTCTGCATGTCAGGTCAT 52

Db 1254 CGAGCATCATCTCTGCATGTCAGGTCAT 1283

RESULT 39

ADE10522

ID ADE10522 standard; DNA; 3284 BP.

XX AC ADE10522;

DT 23-JAN-2004 (first entry)

XX DE Minicell associated DNA #46.

XX KW membrane protein transfer; minicell membrane; biological membr  
XX hyperproliferative disorder; cancer; ds.

XX OS Synthetic.

XX PN US2003199089-A1.

XX PD 23-OCT-2003.

XX PF 28-MAY-2002; 2002US-00157318.

XX PR 05-JUN-2001; 2001US-0295566P.

XX PR 25-FEB-2002; 2002US-0359843P.

XX PA (SURB/) SURBER M W.

XX PA (SABH/) SABBADINI R A.

XX PI Surber MW, Sabbadini RA;

XX DR WPI; 2003-852795/79.

XX PT Transferring a membrane protein from a minicell membrane to a  
XX membrane for diagnosing or treating e.g. cancer by allowing th  
XX and biological membrane to remain in contact for a sufficient  
XX the transfer to occur.

XX PS Disclosure; SEQ ID NO 186; 243pp; English.

ion relates to a method of transferring a membrane protein from a membrane to a biological membrane which comprises contacting a membrane to remain in contact for a period of time sufficient to transfer a protein from a minicell membrane to a biological membrane for a composition for diagnosing or treating hyperproliferative e.g. cancer. The present sequence represents a minicell DNA.

284 BP; 753 A; 865 C; 927 G; 739 T; 0 U; 0 Other;

Similarity 36.9%; Score 28.4; DB 9; Length 3284;

Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CTAGCATCATCCTCTGTCATGTCAGGTCAT 52  
 |||||  
 GAGCATCATCCTCTGTCATGTCAGGTCAT 1283

tandard; DNA; 3284 BP.

4 (first entry)

-2 adrenergic receptor, SalI-XhoI fragment.

eukaryotic expression sequence; open reading frame; ORF;  
 1 minicell; poroplast; spheroplast; protoplast;  
 al cell; anucleate cell; drug discovery; ds; human.

ns.

79-A1.

03.

12; 2002US-00157391.

11; 2001US-0293566P.

12; 2002US-0359843P.

12; 2002US-00154951.

ABADINI R A.

ARKLEY N.

RA, Berkley N;

874920/81.

ug the rate of transfer of nucleic acid from a minicell to a  
 vice versa) useful in the production of achromosomal and  
 cells used for diagnostic and therapeutic applications.

2; SEQ ID NO 186; 242pp; English.

ion relates to determining the rate of transfer of nucleic acid  
 icell to a cell, determining the amount of a nucleic acid  
 ad to a cell from a minicell and detecting the expression of an  
 element in a cell. The minicell comprises an expression  
 ailing eukaryotic expression sequences operably linked to an open  
 frame (ORF) encoding a detectable polypeptide, the minicells  
 binding group and the binding group displays an epitope of the  
 minicell is a eubacterial minicell, a poroplast, a spheroplast  
 oplast. The cell is a eukaryotic cell. The binding group is an  
 or antibody derivative, especially a single-chain antibody, an  
 r an organic compound. The detectable polypeptide is a  
 at polypeptide. The methods are used in the production of

CC achromosomal and anucleate cells useful for applications such  
 CC diagnostic and therapeutic uses, as well as research tools and  
 CC drug discovery. The present sequence is a human DNA sequence u  
 CC minicell construct of the invention. Note: The authors have mi  
 CC SEQ ID numbers between the text and the sequence listing such  
 CC of the sequences cannot be conclusively identified.

XX Sequence 3284 BP; 753 A; 865 C; 927 G; 739 T; 0 U; 0 Other;

Query Match 36.9%; Score 28.4; DB 9; Length 3284;  
 Best Local Similarity 96.7%; Pred. No. 2;  
 Matches 29; Conservative 0; Mismatches 1; Indels 0;

QY 23 CTAGCATCATCCTCTGTCATGTCAGGTCAT 52

Db 1254 CGAGCATCATCCTCTGTCATGTCAGGTCAT 1283

Search completed: April 13, 2004, 18:34:55  
 Job time : 220.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

cleic search, using sw model

April 13, 2004, 17:11:06 ; Search time 1107.5 Seconds  
(without alignments)  
3013.466 Million cell updates/sec

US-10-090-326-25

1 acttacnagccagcancanca.....ggagnacaacagccancagt 77

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

3470272 seqs, 21671516995 residues

hits satisfying chosen parameters: 6940544

length: 0

length: 20000000000

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

GenEmbl:

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_on:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_or:\*

21: em\_ov:\*

22: em\_pat:\*

23: em\_ph:\*

24: em\_pl:\*

25: em\_ro:\*

26: em\_sts:\*

27: em\_un:\*

28: em\_vi:\*

29: em\_htg\_hum:\*

30: em\_htg\_inv:\*

31: em\_htg\_other:\*

32: em\_htg\_mus:\*

33: em\_htg\_pln:\*

34: em\_htg\_rod:\*

35: em\_htg\_man:\*

36: em\_htg\_vrt:\*

37: em\_sy:\*

38: em\_htgo\_hum:\*

39: em\_htgo\_mus:\*

40: em\_htgo\_pln:\*

41: em\_htgo\_rod:\*

), is the number of results predicted by chance to have a

score greater than or equal to the score of the result being  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Descript
1	30	39.0	1059	6	AR061745	AR061745
2	30	39.0	1059	6	AR061746	AR061746
3	30	39.0	1059	6	AR061986	AR061986
4	30	39.0	1059	6	AR061987	AR061987
5	30	39.0	1059	6	AR206220	AR206220
6	30	39.0	1059	6	AR206221	AR206221
7	30	39.0	1452	9	BC027179	BC027179
8	30	39.0	1587	6	AR061747	AR061747
9	30	39.0	1587	6	AR061748	AR061748
10	30	39.0	1587	6	AR061988	AR061988
11	30	39.0	1587	6	AR061989	AR061989
12	30	39.0	1587	6	B59483	B59483
13	30	39.0	1587	6	AR206222	AR206222
14	30	39.0	1587	6	AR206223	AR206223
15	30	39.0	1590	6	AX113825	AX113825
16	30	39.0	1590	9	HSU01873	HSU01873
17	30	39.0	1880	9	HUMTYR	HUMTYR
18	30	39.0	1886	6	I00224	I00224
19	30	39.0	1888	9	HSTYROR	HSTYROR
20	30	39.0	1894	6	AR004653	AR004653
21	30	39.0	1894	6	AR062126	AR062126
22	30	39.0	1894	6	AR062127	AR062127
23	30	39.0	1894	6	I21255	I21255
24	30	39.0	1906	6	AR003571	AR003571
25	30	39.0	1906	6	I17316	I17316
26	30	39.0	1910	6	AR061264	AR061264
27	30	39.0	1929	9	HUMTYRM	HUMTYRM
28	30	39.0	2041	6	AX285154	AX285154
29	30	39.0	2384	6	AR167367	AR167367
30	30	39.0	2384	6	E11466	E11466
31	30	39.0	2384	6	I49609	I49609
32	30	39.0	2384	9	HUMTYRA	HUMTYRA
33	28.8	37.4	155690	10	AC121885	AC121885
34	28.6	37.1	162383	2	AC118220	AC118220
35	28.4	36.9	67	6	BD131906	BD131906
36	28.4	36.9	67	6	BD131908	BD131908
37	28.4	36.9	299	4	AB128155	AB128155
38	28.4	36.9	1063	1	ECO487615	ECO487615
39	28.4	36.9	1063	1	ECO487616	ECO487616
40	28.4	36.9	1063	1	ECO487617	ECO487617
41	28.4	36.9	1950	6	I00682	I00682
42	28.4	36.9	3078	1	ECLACZ	ECLACZ
43	28.4	36.9	3096	6	AR387286	AR387286
44	28.4	36.9	3144	6	AX348046	AX348046
45	28.4	36.9	3144	6	AX353913	AX353913

## ALIGNMENTS

RESULT 1  
AR061745  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES

AR061745  
Sequence 58 from patent US 5843654.  
AR061745  
AR061745.1 GI:5989436  
Unknown.  
Unknown.  
Unclassified.  
1 (bases 1 to 1059)  
Heisler,L.M., Fors,L. and Brow,M.Ann.D.  
Rapid detection of mutations in the p53 gene  
Patent: US 5843654-A 58 01-DEC-1998;  
Location/Qualifiers

AR061745  
1059 bp  
DNA  
linear  
PA

1..1059  
/organism="unknown"  
/mol\_type="unassigned DNA"

39.0%; Score 30; DB 6; Length 1059;  
Similarity 67.9%; Pred. No. 1.9;  
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

ACTTACNAGCCGAGCCAGCAGCATTCTTAGCATCCTCTGCGATG-GTCAGGTCAATTGGAGG 59  
|||||  
ACTTACTCAGCCGAGCAGCATTCTTCTCTGCGAGATTCTCTGTAGCCGATTGGAGG 534

AGNACAAGCCAGCCAGCAGT 77  
|||||  
AGTACAAGCCAGCAGTCACT 552

61746 1059 bp DNA linear PAT 29-SEP-1999  
Sequence 59 from patent US 5843654.  
61746  
61746.1 GI:5989437

nown.  
lassified.  
(bases 1 to 1059)  
ler,L.M., Fors,L. and Brow,M. Ann.D.  
id detection of mutations in the p53 gene  
ent: US 5843654-A 59 01-DEC-1998;  
Location/Qualifiers  
1..1059

/organism="unknown"  
/mol\_type="unassigned DNA"

39.0%; Score 30; DB 6; Length 1059;  
Similarity 67.9%; Pred. No. 1.9;  
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

TTTACNAGCCGAGCCAGCAGCATTCTTAGCATCCTCTGCGATG-GTCAGGTCAATTGGAGG 59  
|||||  
TTTACTCAGCCGAGCAGCATTCTTCTCTGCGAGATTCTCTGTAGCCGATTGGAGG 534

3NACAAGCCAGCCAGCAGT 77  
|||||  
3TACAGCCAGCAGTCACT 552

1986 1059 bp DNA linear PAT 29-SEP-1999  
Sequence 58 from patent US 5843669.  
1986  
1986.1 GI:5989677

nown.  
lassified.  
(bases 1 to 1059)  
er,M.W., Lyamichev,V.I. and Lyamichev,N.  
avage of nucleic acid using thermostable methanococcus  
aschii FEN-1 endonucleases  
nt: US 5843669-A 58 01-DEC-1998;  
Location/Qualifiers  
1..1059

/organism="unknown"  
/mol\_type="unassigned DNA"

39.0%; Score 30; DB 6; Length 1059;

Best Local Similarity 67.9%; Pred. No. 1.9;  
Matches 53; Conservative 0; Mismatches 24; Indels 1

QY 1 ACTTACNAGCCGAGCCAGCAGCATTCTTAGCATCCTCTGCGATG-GTCAGGTCA  
|||||  
Db 475 ACTTACTCAGCCGAGCAGCATTCTTCTCTGCGAGATTCTCTGTAGCCG  
|||||  
QY 60 AGNACAAGCCAGCCAGCAGT 77  
|||||  
Db 535 AGTACAAGCCAGCAGTCACT 552

RESULT 4  
AR061987  
LOCUS AR061987 1059 bp DNA linear PA  
DEFINITION Sequence 59 from patent US 5843669.  
ACCESSION AR061987  
VERSION AR061987.1 GI:5989678  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1059)  
AUTHORS Kaiser,M.W., Lyamichev,V.I. and Lyamichev,N.  
TITLE Cleavage of nucleic acid using thermostable metho  
JOURNAL jannaschii FEN-1 endonucleases  
FEATURES Patent: US 5843669-A 59 01-DEC-1998;  
Location/Qualifiers  
source 1..1059  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 39.0%; Score 30; DB 6; Length 1059;  
Best Local Similarity 67.9%; Pred. No. 1.9;  
Matches 53; Conservative 0; Mismatches 24; Indels 1;

QY 1 ACTTACNAGCCGAGCCAGCAGCATTCTTAGCATCCTCTGCGATG-GTCAGGTCA  
|||||  
Db 475 ACTTACTCAGCCGAGCAGCATTCTTCTCTGCGAGATTCTCTGTAGCCG  
|||||  
QY 60 AGNACAAGCCAGCCAGCAGT 77  
|||||  
Db 535 AGTACAAGCCAGCAGTCACT 552

RESULT 5  
AR206220  
LOCUS AR206220 1059 bp DNA linear PAT  
DEFINITION Sequence 58 from patent US 6372424.  
ACCESSION AR206220  
VERSION AR206220.1 GI:21504758  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1059)  
AUTHORS Brow,M. Ann.D., Lyamichev,V.I. and Olive,D. Michael.  
TITLE Rapid detection and identification of pathogens  
JOURNAL Patent: US 6372424-A 58 16-APR-2002;  
FEATURES Location/Qualifiers  
source 1..1059  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 39.0%; Score 30; DB 6; Length 1059;  
Best Local Similarity 67.9%; Pred. No. 1.9;  
Matches 53; Conservative 0; Mismatches 24; Indels 1;

QY 1 ACTTACNAGCCGAGCCAGCAGCATTCTTAGCATCCTCTGCGATG-GTCAGGTCA  
|||||  
Db 475 ACTTACTCAGCCGAGCAGCATTCTTCTCTGCGAGATTCTCTGTAGCCGAT

o sapiens (human)  
o sapiens  
ayota; Metazoa; Craniata; Craniata; Vertebrata; Euteleostomi;  
malia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
(bases 1 to 1452)  
ausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
usner, R.D., Collins, F.S., Wagner, L., Sherman, C.M., Schuler, G.D.,  
schul, S.F., Zeeberg, H., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
kins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
tchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
pleton, M., Soares, M.B., Bonaldo, M.F., Casavani, T.H.,  
eetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
ninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
amson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,  
ernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
ley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
lalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
ey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,  
chez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
ffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
kson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
terfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E.,  
nerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.,  
eration and initial analysis of more than 15,000 full-length  
an and mouse cdna sequences  
c. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

GNACACAGCCATCAGT 77  
|||||  
GTACACAGCCATCAGT 920

61747 1587 bp DNA linear PAT 29-SEP-1999  
ence 60 from patent US 5843654.  
61747  
61747.1 GI:5989438

nown.  
nown.  
classified.  
(bases 1 to 1587)  
sler, L.M., Fors, L. and Brow, M. Ann. D.  
id detection of mutations in the p53 gene  
ent: US 5843654-A 60 01-DEC-1998;  
Location/Qualifiers  
1. .1587  
/organism="unknown"  
/mol\_type="unassigned DNA"

39.0%; Score 30; DB 6; Length 1587;  
milarity 67.9%; Pred. No. 1.9;  
Conservative 0; Mismatches 24; Indels 1; Gaps 1;  
ACTTACNAGCCAGCCATCAGT 77  
|||||  
ACTTACTCAGCCAGCCATCAGT 859

AGNACACAGCCATCAGT 77  
|||||  
GTACACAGCCATCAGT 859

61748 1587 bp DNA linear PAT 29-SEP-1999  
ence 61 from patent US 5843654.  
61748  
61748.1 GI:5989439

nown.  
known.  
classified.  
(bases 1 to 1587)  
sler, L.M., Fors, L. and Brow, M. Ann. D.  
id detection of mutations in the p53 gene  
ent: US 5843654-A 61 01-DEC-1998;  
Location/Qualifiers  
1. .1587  
/organism="unknown"  
/mol\_type="unassigned DNA"

39.0%; Score 30; DB 6; Length 1587;  
milarity 67.9%; Pred. No. 1.9;  
Conservative 0; Mismatches 24; Indels 1; Gaps 1;  
ACTTACNAGCCAGCCATCAGT 77  
|||||  
ACTTACTCAGCCAGCCATCAGT 859

AGNACACAGCCATCAGT 77  
|||||  
AGTACACAGCCATCAGT 859

LOCUS AR061988 1587 bp DNA linear PAT  
DEFINITION Sequence 60 from patent US 5843669.  
ACCESSION AR061988  
VERSION AR061988.1 GI:5989679  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1587)  
AUTHORS Kaiser, M.W., Lyamichev, V.I. and Lyamichev, N.  
TITLE Cleavage of nucleic acid using thermostable metho:  
Jannaschii FEN-1 endonucleases  
JOURNAL Patent: US 5843669-A 60 01-DEC-1998;  
FEATURES Location/Qualifiers  
source 1. .1587  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 39.0%; Score 30; DB 6; Length 1587;  
Best Local Similarity 67.9%; Pred. No. 1.9;  
Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNAGCCAGCCATCAGT 77  
|||||  
Db 782 ACTTACTCAGCCAGCCATCAGT 859  
Qy 60 AGNACACAGCCATCAGT 77  
|||||  
Db 842 AGTACACAGCCATCAGT 859

RESULT 11  
AR061989 1587 bp DNA linear PA  
LOCUS AR061989  
DEFINITION Sequence 61 from patent US 5843669.  
ACCESSION AR061989  
VERSION AR061989.1 GI:5989680  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1587)  
AUTHORS Kaiser, M.W., Lyamichev, V.I. and Lyamichev, N.  
TITLE Cleavage of nucleic acid using thermostable metho:  
Jannaschii FEN-1 endonucleases  
JOURNAL Patent: US 5843669-A 61 01-DEC-1998;  
FEATURES Location/Qualifiers  
source 1. .1587  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 39.0%; Score 30; DB 6; Length 1587;  
Best Local Similarity 67.9%; Pred. No. 1.9;  
Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNAGCCAGCCATCAGT 77  
|||||  
Db 782 ACTTACTCAGCCAGCCATCAGT 859  
Qy 60 AGNACACAGCCATCAGT 77  
|||||  
Db 842 AGTACACAGCCATCAGT 859

RESULT 12  
E59483 1587 bp DNA linear PA  
LOCUS E59483  
DEFINITION Method for determining effect of gray hair-preventing ag  
ACCESSION E59483  
VERSION E59483.1 GI:18622554  
KEYWORDS JP 2000300298-A/3.

no sapiens (human)  
 no sapiens  
 Caryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 (bases 1 to 1587)  
 tat.t. and Ifuku, O.  
 :hod for determining effect of gray hair-preventing agent  
 :ent: JP 2000300298-A 3 31-OCT-2000;  
 :SEIDO CO LTD  
 Homo sapiens (human)  
 JP 2000300298-A/3  
 31-OCT-2000  
 23-APR-1999 JP 1999117006  
 TATSURO IDETA,OJI IFUKU  
 C12Q1/68,C12N15/09,G01N33/15,G01N33/50,C12N15/00 CC  
 Key Location/Qualifiers  
 source 1..1587  
 /organism="Homo sapiens (human)"  
 Location/Qualifiers  
 1..1587  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 39.0%; Score 30; DB 6; Length 1587;  
 milarity 67.9%; Pred. No. 1.9; Indels 1; Gaps 1;  
 Conservative 0; Mismatches 24; Indels 1; Gaps 1;  
 CTTACNAGCCAGCCAGCATTCTAGCATCATCTCTGCGATG-GTCAGGTCATTGGAGG 59  
 |||||  
 CTTACTCAGCCAGCATCATCTCTCTCTTGGCAGATTGCTGTAGCCGATTGGAGG 841  
 |||||  
 GNACACAGCCAGCAGT 77  
 |||||  
 GTACACAGCCATCAGT 859  
 |||||  
 06222 1587 bp DNA linear PAT 20-JUN-2002  
 uence 60 from patent US 6372424.  
 06222  
 06222.1 GI:21504760  
 nown.  
 nown.  
 lassified.  
 (bases 1 to 1587)  
 w.M. Ann.D., Lyamichev, V.I. and Olive, D. Michael.  
 id detection and identification of pathogens  
 ent: US 6372424-A 60 16-APR-2002;  
 Location/Qualifiers  
 1..1587  
 /organism="unknown"  
 /mol\_type="unassigned DNA"  
 39.0%; Score 30; DB 6; Length 1587;  
 milarity 67.9%; Pred. No. 1.9; Indels 1; Gaps 1;  
 Conservative 0; Mismatches 24; Indels 1; Gaps 1;  
 CTTACNAGCCAGCCAGCATTCTAGCATCATCTCTGCGATG-GTCAGGTCATTGGAGG 59  
 |||||  
 CTTACTCAGCCAGCATCATCTCTCTCTTGGCAGATTGCTGTAGCCGATTGGAGG 841  
 |||||  
 GNACACAGCCAGCAGT 77  
 |||||  
 GTACACAGCCATCAGT 859  
 |||||

AR206223  
 LOCUS 1587 bp DNA linear PA  
 DEFINITION Sequence 61 from patent US 6372424.  
 ACCESSION AR206223  
 VERSION AR206223.1 GI:21504761  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 Unclassified.  
 REFERENCE 1 (bases 1 to 1587)  
 AUTHORS Brow, M. Ann. D., Lyamichev, V. I. and Olive, D. Michael.  
 TITLE Rapid detection and identification of pathogens  
 JOURNAL Patent: US 6372424-A 61 16-APR-2002;  
 FEATURES Location/Qualifiers  
 source 1..1587  
 /organism="unknown"  
 /mol\_type="unassigned DNA"  
 ORIGIN  
 Query Match 39.0%; Score 30; DB 6; Length 1587;  
 Best Local Similarity 67.9%; Pred. No. 1.9;  
 Matches 53; Conservative 0; Mismatches 24; Indels 1;  
 QY 1 ACTTACNAGCCAGCCAGCATTCTAGCATCATCTCTGCGATG-GTCAGGTCAG 77  
 |||||  
 Db 782 ACTTACTCAGCCAGCATCATCTCTCTCTTGGCAGATTGCTGTAGCCG 77  
 |||||  
 QY 60 AGNACACAGCCAGCAGT 77  
 |||||  
 Db 842 AGTACACAGCCATCAGT 859  
 |||||  
 RESULT 15  
 AX113825 1590 bp DNA linear PAJ  
 LOCUS  
 DEFINITION Sequence 1 from Patent WO0127295.  
 ACCESSION AX113825  
 VERSION AX113825.1 GI:13939991  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hc  
 1  
 Schandendorf, D., Paschen, A., Chakraborty, T. and Domann, E  
 Recombinant attenuated listerias for immunotherapy  
 Patent: WO 0127295-A 1 19-APR-2001;  
 Deutsches Krebsforschungszentrum Stiftung des Oeffentli  
 (DE)  
 Location/Qualifiers  
 1..1590  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 1..1590  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC37863.1"  
 /db\_xref="GI:13939992"  
 /db\_xref="REMBL:CAC37863"  
 /translation="MLLAVLYCLLWSFQTSAGHFPFRACVSKNLMX  
 PCGUSGRGSCQNLISNAPLGQFPFTGVDRESWPSVFNRTQ  
 CKFGWPNCTERRLLVRNIFDLISAPKDKFFAYLTKAKHTISSD  
 NGSTPMFNDINIDLVFVMMHYVSMALGGVEIWRDIDFAHEAPAI  
 EGEIKLTGDNFTIPYWDWDKAEKDCIDCTDEYMGCHPTNPILLS  
 SLELYNSHQSLCNGTPEGPLRRNPGNHDKXRTPLRPSADVEFCL  
 RANFSRNTLEGFASFLGIADASQSMNHAIHYMNGTMSQVQS  
 VDSIFEQWLRHRPLQYVPEANAFIGNRESYVMPFPLFNGDF  
 YLQSDPDSFDQYTKYLEQASRIWSMLLGAAMVGAVALTALLAGLV  
 BEKQLLMKEEDYHSLYQSHL"  
 FEATURES  
 source  
 CDS  
 ORIGIN  
 Query Match 39.0%; Score 30; DB 6; Length 1590;

imilarity 67.9%; Pred. No. 1.9; Mismatches 0; Indels 24; Gaps 1;  
 ; Conservative 0; Mismatches 24; Indels 1; Gaps 1;  
 ACTTACNAGCCAGCCAGCANCATTCTAGCATCCTCTGCGATG-GTCAGGTCAATTTGGAGG 59  
 |||||  
 ACTTACTAGCCAGCATCATTCTTCTCTCTGCGAGATTCTCTAGCCGATTGGAGG 841  
 |||||  
 AGNACAGCAGCCAGCAGT 77  
 |||||  
 AGTACACAGCCATCAGT 859  
 |||||

U01873 1590 bp mRNA linear PRI 27-SEP-1993  
 Human tyrosinase mRNA, complete cds.  
 1873  
 1873.1 GI:403421  
 no sapiens (human)  
 no sapiens  
 Karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 (bases 1 to 1590)  
 ichard,V., Van Pel,A., Wolfel,T., Wolfel,C., De Plaen,E.,  
 che,B., Coulie,P. and Boon,T.  
 a tyrosinase gene codes for an antigen recognized by autologous  
 T lymphocytes on HLA-A2 melanomas  
 Exp. Med. 178 (2), 489-495 (1993)  
 340625  
 10755  
 (bases 1 to 1590)  
 ichard,B., Fuller,B.B., Vijayasaradhi,S. and Houghton,A.N.  
 unction of pigmentation in mouse fibroblasts by expression of  
 human tyrosinase cDNA  
 Exp. Med. 169 (6), 2029-2042 (1989)  
 379151  
 99655  
 (bases 1 to 1590)  
 on,B.S., Haq,A.K., Pomerantz,S.H. and Halaban,R.  
 ation and sequence of a cDNA clone for human tyrosinase that  
 s at the mouse c-albino locus  
 oc. Natl. Acad. Sci. U.S.A. 84 (21), 7473-7477 (1987)  
 241128  
 23263  
 (bases 1 to 1590)  
 ichard,V.G.  
 ect Submission  
 mitted (16-SEP-1993) Vincent G Brichard, Ludwig Institute for  
 icer Research, Avenue Hippocrate, 74, Brussels B-1200 Belgium  
 Location/Qualifiers  
 1. 1590  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 1. 1590  
 /codon\_start=1  
 /product="tyrosinase"  
 /protein\_id="AAB60319.1"  
 /db\_xref="GI:403422"  
 /translation="MLKAVLYCLLWSFQTSAGHPFACVSSKNLMEKECCPPWGGDRS  
 PCQLSGRGSCQNTLLSNAPLGQFPFTGVDRESWPSVFNRTCCQSGNPMFGNCGN  
 CKFGFPGNCTERRLLVRNFIPLSAPEKDFAYLTAKHTISDDYVPIGTYGQMK  
 NGSTPMFNDINIYDLFVMMHYVVSMDALLGSEIWRDIDFAHEAPAFIPWRLLFRW  
 EQETKLTGDSNFPIPYWMDRAEKDCICTDEYGGQHTPNLNLSPASFPSSQIVC  
 SRLEYNHQSCLNGTPEGPRPNPNHDKSRTPRLPSADVEFCLSLTQYESGMDK  
 AANFSRNTLGGFASPLTGIAQSSSMHNAHLYMNGTMSQVQSANDPIFLHAF  
 VDSIFQWLQRRLQEVPEANAPIGHNRESYVMPFPLVRNGDFFTSKDLGVYS  
 YLQSDPDSFDQYIKSYLEQASRIWSLLGAAMGAVLTALLAGIVSLLCRRKKQLP  
 EEKQPLMEKEDYHSLYQSHL"

39.0%; Score 30; DB 9; Length 1590;

Best Local Similarity 67.9%; Pred. No. 1.9; Mismatches 0; Indels 1;  
 Matches 53; Conservative 0; Mismatches 24; Indels 1  
 QY 1 ACTTACNAGCCAGCCAGCANCATTCTAGCATCCTCTGCGATG-GTCAGGTCA  
 |||||  
 Db 782 ACTTACTAGCCAGCATCATTCTTCTCTCTGCGAGATTCTCTGCGAGATTCTCTGAGCCG  
 |||||  
 QY 60 AGNACAGCAGCCAGCAGT 77  
 |||||  
 Db 842 AGTACACAGCCATCAGT 859  
 |||||

RESULT 17  
 HUMTYR  
 LOCUS  
 DEFINITION Human tyrosinase mRNA, 3' end.  
 ACCESSION J03581  
 VERSION J03581.1 GI:340027  
 KEYWORDS tyrosinase.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 REFERENCE 1 (bases 1 to 1880)  
 AUTHORS Kwon,B.S., Haq,A.K., Pomerantz,S.H. and Halaban,R.  
 TITLE Isolation and sequence of a cDNA clone for human tyros  
 maps at the mouse c-albino locus  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84 (21), 7473-7477 (1987)  
 MEDLINE 88041128  
 PUBMED 2823263  
 REFERENCE 2 (bases 1144 to 1190)  
 AUTHORS Kwon,B.S., Haq,A.K., Pomerantz,S.H. and Halaban,R.  
 TITLE Correction: Isolation and sequence of a cDNA clone for  
 tyrosinase that maps at the mouse c-albino locus  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85, 6352-6352 (1988)  
 COMMENT Original source text: Human melanocyte, cDNA to mRNA,  
 [2] revises [1].  
 Draft entry and printed copy of sequence for [1] kind:  
 B.S.Kwon, 29-SEP-1987.  
 Location/Qualifiers  
 1. 1880  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /map="11q14-q21"  
 1. 1880  
 /gene="TYR"  
 <1. 1880  
 /gene="TYR"  
 /product="tyrosinase mRNA"  
 <1. 1883  
 /gene="TYR"  
 /note="tyrosinase precursor (EC 1.14.18.1)"  
 /codon\_start=1  
 /protein\_id="AAA61241.1"  
 /db\_xref="GI:340028"  
 /db\_xref="GDB:G00-120-476"  
 /translation="LLLAIVLYCLLWSFQTSAGHPFACVSSKNLMI  
 CQLSGRGSCQNTLLSNAPLGQFPFTGVDRESWPSVFNRTCCQ  
 KFGFPGNCTERRLLVRNFIPLSAPEKDFAYLTAKHTISDDY  
 GSTPMFNDINIYDLFVMMHYVVSMDALLGSEIWRDIDFAHEAPAI  
 QETKLTGDSNFPIPYWMDRAEKDCICTDEYGGQHTPNLNLISH  
 RLEYNHQSCLNGTPEGPRPNPNHDKSRTPRLPSADVEFCLS  
 ANFSRNTLGGFASPLTGIAQSSSMHNAHLYMNGHVEPTGSAL  
 SIFEQWLQRRLQEVPEANAPIGHNRESYVMPFPLVRNGDFF  
 QSDPDSFDQYIKSYLEQASRIWSLLGAAMGAVLTALLAGVPSI  
 KQPLMEKEGLPLQVSEPFPIKGLGNRVGPKSPDLTLTQSNVQVPE  
 <1. 36  
 /gene="TYR"  
 /note="tyrosinase signal peptide"  
 37. 1680  
 /gene="TYR"  
 /product="tyrosinase"

sig\_peptide  
 mat\_peptide

Merger Center, 1250 First Avenue, New York, N.Y. 10028, USA

Query Match 39.0%: Score 30: DB 6: Length 1894:

imilarity 67.9%; Pred. No. 1.9;  
 ; Conservative 0; Mismatches 24; Indels 1; Gaps 1;  
 ACTTACNAGCCGAGCAGCAGCATCTTCTGATG-GTCAGTCAATTGGAGG 59  
 |||||  
 ACTTACTCAGCCGAGCAGCATCTTCTCTCTTGGCAGATTGCTGTAGCCGATTGGAGG 847  
 |||||  
 AGNACAACAGCCAGCAGT 77  
 |||||  
 AGTACAACAGCCATCAGT 865  
 |||||

062126 1894 bp DNA linear PAT 29-SEP-1999  
 quence 1 from patent US 5843687.  
 062126  
 062126.1 GI:5989817  
 known.  
 known.  
 classified.  
 (bases 1 to 1894)  
 lfel,T., Pel,A.Van., Brichard,V. and Boon-Falleur,T.  
 olated, tyrosinase derived peptides and uses thereof  
 tent: US 5843687-A 1 01-DEC-1998;  
 Location/Qualifiers  
 1..1894  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

imilarity 39.0%; Score 30; DB 6; Length 1894;  
 ; Conservative 0; Mismatches 24; Indels 1; Gaps 1;  
 ACTTACNAGCCGAGCAGCAGCATCTTCTGATG-GTCAGTCAATTGGAGG 59  
 |||||  
 ACTTACTCAGCCGAGCAGCATCTTCTCTCTTGGCAGATTGCTGTAGCCGATTGGAGG 847  
 |||||  
 AGNACAACAGCCAGCAGT 77  
 |||||  
 AGTACAACAGCCATCAGT 865  
 |||||

362127 1894 bp DNA linear PAT 29-SEP-1999  
 quence 1 from patent US 5843688.  
 362127  
 362127.1 GI:5989818  
 known.  
 known.  
 classified.  
 (bases 1 to 1894)  
 lfel,T., Van Pel,A., Brichard,V., Boon-Falleur,T., Deplaen,E.,  
 lie,P., Renauld,J.-C. and Lethe,B.  
 plated tyrosinase derived peptides and uses thereof  
 tent: US 5843688-A 1 01-DEC-1998;  
 Location/Qualifiers  
 1..1894  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

imilarity 39.0%; Score 30; DB 6; Length 1894;  
 ; Conservative 0; Mismatches 24; Indels 1; Gaps 1;  
 ACTTACNAGCCGAGCAGCAGCATCTTCTGATG-GTCAGTCAATTGGAGG 59  
 |||||  
 ACTTACTCAGCCGAGCAGCATCTTCTCTCTTGGCAGATTGCTGTAGCCGATTGGAGG 847  
 |||||

QY 60 AGNACAACAGCCAGCAGT 77  
 |||||  
 Db 848 AGTACAACAGCCATCAGT 865  
 |||||

RESULT 23  
 LOCUS 121255 1894 bp DNA linear PA  
 DEFINITION Sequence 1 from patent US 5519117.  
 ACCESSION 121255  
 VERSION 121255.1 GI:1601609  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 UNCLASSIFIED.  
 REFERENCE 1 (bases 1 to 1894)  
 AUTHORS Wolfel,T., Pel,A.V., Brichard,V. and Boon-Falleur,T.  
 TITLE Isolated, tyrosinase derived peptides and uses thereof  
 JOURNAL Patent: US 5519117-A 1 21-MAY-1996;  
 FEATURES Location/Qualifiers  
 . source  
 1..1894  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

Query Match 39.0%; Score 30; DB 6; Length 1894;  
 Best Local Similarity 67.9%; Pred. No. 1.9;  
 Matches 53; Conservative 0; Mismatches 24; Indels 1

QY 1 ACTTACNAGCCGAGCAGCAGCATCTTCTGATG-GTCAGTCA  
 |||||  
 Db 788 ACTTACTCAGCCGAGCAGCATCTTCTCTCTTGGCAGATTGCTGTAGCCG  
 |||||

QY 60 AGNACAACAGCCAGCAGT 77  
 |||||  
 Db 848 AGTACAACAGCCATCAGT 865  
 |||||

RESULT 24  
 LOCUS AR003571 1906 bp DNA linear PA  
 DEFINITION Sequence 1 from patent US 5744316.  
 ACCESSION AR003571  
 VERSION AR003571.1 GI:3964830  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 UNCLASSIFIED.  
 REFERENCE 1 (bases 1 to 1906)  
 AUTHORS Lethe,B., Brichard,V., Van Pel,A., Wolfel,T. and Boon-  
 TITLE Isolated, tyrosinase derived peptides and uses thereof  
 JOURNAL Patent: US 5744316-A 1 28-APR-1998;  
 FEATURES Location/Qualifiers  
 . source  
 1..1906  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

Query Match 39.0%; Score 30; DB 6; Length 1906;  
 Best Local Similarity 67.9%; Pred. No. 1.9;  
 Matches 53; Conservative 0; Mismatches 24; Indels 1

QY 1 ACTTACNAGCCGAGCAGCAGCATCTTCTGATG-GTCAGTCA  
 |||||  
 Db 806 ACTTACTCAGCCGAGCAGCATCTTCTCTCTTGGCAGATTGCTGTAGCCG  
 |||||

QY 60 AGNACAACAGCCAGCAGT 77  
 |||||  
 Db 866 AGTACAACAGCCATCAGT 883  
 |||||

RESULT 25





llie,J., Brown,J.L., Bolt,A. and van Huffel,C.  
vel genes, compositions and methods for the identification,  
ssessment, prevention, and therapy of human cancers  
tent: WO 0179556-A 959 25-OCT-2001;  
llennium Predictive Medicine, Inc. (US)

Location/Qualifiers  
1..2041  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

imilarity 39.0%; Score 30; DB 6; Length 2041;

; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

ACTTACNAGCCGAGCAGCAGCATCTAGCATCATCTCTGCATG-GTCAGGTCAATTGGAGG 59  
|||||  
ACTTACTAGCCGAGCAGCATCTCTCTCTTGGCAGATTCTGTAGCCGATTGGAGG 966

AGNACACAGCCAGCAGCAGT 77

|||||  
AGTACACAGCCATCAGT 984

167367 2384 bp DNA linear PAT 17-DEC-2001  
quence 3 from patent US 6287569.

167367

167367.1 GI:17903142

known.

classified.  
(bases 1 to 2384)

pps,T.J. and Wu,Y.

scines with enhanced intracellular processing

ent: US 6287569-A 3 11-SEP-2001;

Location/Qualifiers

1..2384

/organism="unknown"

/mol\_type="unassigned DNA"

imilarity 39.0%; Score 30; DB 6; Length 2384;

; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

CTTACNAGCCGAGCAGCAGCATCTAGCATCATCTCTGCATG-GTCAGGTCAATTGGAGG 59

|||||

CTTACTAGCCGAGCAGCATCTCTCTCTTGGCAGATTCTGTAGCCGATTGGAGG 1343

AGNACACAGCCAGCAGT 77

|||||

GTACACAGCCATCAGT 1361

.466 2384 bp RNA linear PAT 29-SEP-1997

IA encoding tyrosinase.

.466

.466.1 GI:22025102

1936140699-A/1.

to sapiens (human)

to sapiens

aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

(bases 1 to 2384)

.bata,T., Suzuki,S., Takimoto,H. and Masui,S.

ASUREMENT OF TYROSINASE MESSENGER RNA AMOUNT

JOURNAL Patent: JP 1996140699-A 1 04-JUN-1996;

COMMENT POLA CHEM IND INC

OS Homo sapiens (human)

PN JP 1996140699-A/1

PD 04-JUN-1996

PF 22-NOV-1994 JP 1994288041

PI SHIBATA TAKASHI, SUZUKI SATOSHI, TAKIMOTO HIROYUK

SHIGEKI

PC C12Q1/68,C12N15/09;

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

FH Key Location/Qualifiers

FH source

1..2384

/organism="Homo sapiens"

/tissue\_type="Skin"

/cell\_type="melanocyte"

1..502

5'UTR

CDS

503..2092

/product="tyrosinase"

2093..2384.

3'UTR Location/Qualifiers

1..2384

/organism="Homo sapiens"

/mol\_type="genomic RNA"

/db\_xref="taxon:9606"

# ORIGIN

Query Match 39.0%; Score 30; DB 6; Length 2384;  
Best Local Similarity 67.9%; Pred. No. 1.9;  
Matches 53; Conservative 0; Mismatches 24; Indels 1

QY 1 ACTTACNAGCCGAGCAGCAGCATCTAGCATCATCTCTGCATG-GTCAGGTCA

|||||

Db 1284 ACTTACTAGCCGAGCAGCATCTCTCTCTTGGCAGATTGTCTGTAGCCG

QY 60 AGNACACAGCCAGCAGT 77

|||||

Db 1344 AGTACACAGCCATCAGT 1361

# RESULT 31

149609

LOCUS I49609 2384 bp DNA linear PA

DEFINITION Sequence 1 from patent US 5641508.

ACCESSION I49609

VERSION I49609.1 GI:2471829

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2384)

AUTHORS Li,L. and Lishko,V.K.

TITLE Method for delivering melanin to hair follicles

JOURNAL Patent: US 5641508-A 1 24-JUN-1997;

FEATURES Location/Qualifiers

source

1..2384

/organism="unknown"

/mol\_type="unassigned DNA"

# ORIGIN

Query Match 39.0%; Score 30; DB 6; Length 2384;  
Best Local Similarity 67.9%; Pred. No. 1.9;  
Matches 53; Conservative 0; Mismatches 24; Indels 1;

QY 1 ACTTACNAGCCGAGCAGCAGCATCTAGCATCATCTCTGCATG-GTCAGGTCAI

|||||

Db 1284 ACTTACTAGCCGAGCAGCATCTCTCTCTTGGCAGATTGTCTGTAGCCG

QY 60 AGNACACAGCCAGCAGT 77

|||||

Qy 1 ACTTACNAGCCAGCANCATTCTAGCATCTCTGCATG-GTCAGGTCA'

Db 1284 ACTTACTCAGCCAGCATCTCTCTCTTGGCAGATTGTCTGTAGCGI

db 1344 AGTACAACAGCCATCAGT 1361

no sapiens (human)

Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo

Ballard, S., Tomita, Y., Tagami, H., Muller, R.M. and Cohen, T.

22868

Meda, A., Tomita, Y., Okinaga, S., Tagami, H. and Shibahara, S. Functional analysis of the cDNA encoding human tyrosinase precursor. *Biochem. Biophys. Res. Commun.* 162 (3), 984-990 (1989)

Dec 2, 1996 this sequence replaced qi:340036.

Location/Qualifiers

```

/organism="Homo sapiens"
/mol_type="mrna"

```

`/chromosome="11"`  
`/map="111014-01"`

/tissue\_type="pla

1 392. 399

```

/gene="TYR"

```

```
303: .2092  
/gene="TPP"
```

```
/codon_start=1
```

```
/db_xref="GI:3400
```

PCGQLSGRGSCQNILLS

EQEIQKLTGDENFTIPY

HANF SF KNI IPEGFASF  
 VDSIFFEOWI.BRRHBPLOE

EERQPLMEKEDYHSLY  
503 EEC

e 557. 2089

milarity 67.9%; Pred

## SPING INFORMATION:

Sping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see [p://genome.wustl.edu](http://genome.wustl.edu)

## JCE INFORMATION:

a RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57Bl/6J mouse spleen 3/for brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

## IGHBORING SEQUENCE INFORMATION:

is sequence is the entire insert of the clone.

## Location/Qualifiers

1. .155690

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/chromosome="3"

/map="3"

/clone="RP24-138N20"

/clone\_lib="RPCI-24"

221. .366

/rpt\_family="MaLR"

389. .597

/rpt\_family="B2"

612. .791

/rpt\_family="L1"

992. .1278

/rpt\_family="MaLR"

2544. .2646

/rpt\_family="Alu"

3099. .3989

/rpt\_family="RMR13B"

4291. .4371

/rpt\_family="B2"

4380. .4589

/rpt\_family="B2"

4628. .4967

/rpt\_family="ERV1"

6288. .6485

/rpt\_family="B2"

6574. .6885

/rpt\_family="MaLR"

6886. .7362

/rpt\_family="ERV1"

7848. .7942

/rpt\_family="B4"

8007. .8079

/rpt\_family="B4"

9216. .9370

/rpt\_family="B4"

10005. .10075

/rpt\_family="ERVX"

11588. .11859

/rpt\_family="B4"

12238. .12631

/rpt\_family="MaLR"

13019. .13179

/rpt\_family="B4"

13528. .13800

/rpt\_family="B4"

13927. .14064

/rpt\_family="Alu"

16826. .18628

/rpt\_family="L1"

20265. .20685

/rpt\_family="MaLR"

20709. .20884

/rpt\_family="B2"

20916. .21017

/rpt\_family="MaLR"

repeat\_region 22202. .22485  
/rpt\_family="MaLR"  
repeat\_region 23005. .23191  
/rpt\_family="B2"  
repeat\_region 23270. .23333  
/rpt\_family="ERV1"  
repeat\_region 23900. .24102  
/rpt\_family="B2"  
repeat\_region 24278. .24333  
/rpt\_family="B4"  
repeat\_region 24402. .24548  
/rpt\_family="Alu"  
repeat\_region 24588. .24746  
/rpt\_family="B4"  
repeat\_region 28996. .29273  
/rpt\_family="MER2\_type"  
repeat\_region 29274. .29704  
/rpt\_family="RMR17C"  
repeat\_region 29705. .29783  
/rpt\_family="MER2\_type"  
repeat\_region 29903. .29980  
/rpt\_family="MER2\_type"  
repeat\_region 30871. .31011  
/rpt\_family="Alu"  
repeat\_region 32984. .33155  
/rpt\_family="B2"  
repeat\_region 33641. .33787  
/rpt\_family="L1"  
repeat\_region 34591. .34640  
/rpt\_family="B2"  
repeat\_region 34726. .34810  
/rpt\_family="Alu"  
repeat\_region 37550. .37651  
/rpt\_family="Alu"  
repeat\_region 39105. .39246  
/rpt\_family="L1"  
repeat\_region 43851. .43916  
/rpt\_family="L2"  
repeat\_region 44106. .44246  
/rpt\_family="B2"  
repeat\_region 44793. .44986  
/rpt\_family="B2"  
repeat\_region 46081. .46304  
/rpt\_family="MER1\_type"  
repeat\_region 46490. .46712  
/rpt\_family="MIR"  
repeat\_region 47296. .47449  
/rpt\_family="MaLR"  
repeat\_region 47605. .47719  
/rpt\_family="MaLR"  
repeat\_region 48176. .48247  
/rpt\_family="L2"  
repeat\_region 48830. .49151  
/rpt\_family="ERV1"  
repeat\_region 49193. .49579  
/rpt\_family="MaLR"  
repeat\_region 50485. .50687  
/rpt\_family="B2"  
repeat\_region 53587. .53710  
/rpt\_family="Alu"  
repeat\_region 55735. .56206  
/rpt\_family="ERV1"  
repeat\_region 58819. .59027  
/rpt\_family="L1"  
repeat\_region 59050. .59190  
/rpt\_family="L1"

Query Match 37.4%; Score 28.8; DB 10; Length 1556;  
Best Local Similarity 62.7%; Pred.No. 6;  
Matches 42; Conservative 0; Mismatches 25; Indels 0;

Oy 11 OCCAGCNCATTCAGCATCATCTCTGTCATGTCAGTCATTTGGAGGAGN

(bases 1 to 162383)  
ren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
na,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouckgaeter,B.,  
arata,J., Chang,J., Chazaro,B., Choepel,Y.I., Collymore,A.,  
k.A., Cooke,P., DeBrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
O.S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,  
dyna,S., Gord,S., Graham,L., Grand-pierre,N., Hagos,B.,  
ton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
atas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,  
rath,M., Meldrum,J., Meneus,L., Mihoval,T., Miengva,V.,  
phy,T., Taylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
monor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
nkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
an,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
th.C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Tallamas,  
faye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
I.R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
bek,L., Zimmer,A. and Zody,M.

ect Submission

mitted (28-Aug-2002) Whitehead Institute/MIT Center for Genome  
earch, 320 Charles Street, Cambridge, MA 02141, USA

Aug 28, 2002 this sequence version replaced gi:22267708.

Repeats were identified using RepeatMasker:

etic construct  
ical sequences.

```

(bases 1 to 67)
ewart,F., Zhang,Y. and Buchholz,F.
vel DNA cloning method
tent: JP 2002503448-A 24 05-FEB-2002;
ROPAEISCHES LABORATORIUM FUER MOLEKULARBIOLOGIE
Artificial Sequence
JP 2002503448-A/24
05-FEB-2002
07-DEC-1998 JP 2000524410
05-DEC-1997 EP 97121462.2,05-OCT-1998 EP 98118756.0 PI
ANCIS STEWART,YOUNG ZHANG,FRANK BUCHHOLZ
C12N15/09/(C12N15/09,C12R1:19),C12N15/00,C12R1:19)
Description of Artificial Sequence: Artificial Sequence FH
y source Location/Qualifiers
1..67
/organism='Artificial Sequence'.
Location/Qualifiers
1..67
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
36.9%; Score 28.4; DB 6; Length 67;
imilarity 96.7%; Pred. No. 7.4;
; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CTGACATCATCTCTGCATGGTCAGGTGCAT 52
|||||
CGAGCATCATCTCTGCATGGTCAGGTGCAT 32
|||||
131908 67 bp DNA linear PAT 18-SEP-2002
vel DNA cloning method.
131908
131908.1 GI:23226853
2002503448-A/26.
athetic construct
athetic construct
tificial sequences.
(bases 1 to 67)
ewart,F., Zhang,Y. and Buchholz,F.
vel DNA cloning method
tent: JP 2002503448-A 26 05-FEB-2002;
ROPAEISCHES LABORATORIUM FUER MOLEKULARBIOLOGIE
Artificial Sequence
JP 2002503448-A/26
05-FEB-2002
07-DEC-1998 JP 2000524410
05-DEC-1997 EP 97121462.2,05-OCT-1998 EP 98118756.0 PI
ANCIS STEWART,YOUNG ZHANG,FRANK BUCHHOLZ
C12N15/09/(C12N15/09,C12R1:19),C12N15/00,C12R1:19)
Description of Artificial Sequence: Artificial Sequence FH
/ source Location/Qualifiers
1..67
/organism='Artificial Sequence'.
Location/Qualifiers
1..67
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
36.9%; Score 28.4; DB 6; Length 67;
imilarity 96.7%; Pred. No. 7.4;
; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
TAGCATCATCTCTGCATGGTCAGGTGCAT 52
|||||
CGAGCATCATCTCTGCATGGTCAGGTGCAT 40
|||||

```

```

RESULT 37
AB128155
LOCUS
DEFINITION
Oryctolagus cuniculus lac Z mRNA for beta-D-galactosid
cds.
ACCESSION
AB128155.1 GI:39725589
VERSION
AB128155
KEYWORDS
SOURCE
ORGANISM
Oryctolagus cuniculus (rabbit)
Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus
1
Murata,T.
Vascular endothelium has a local anti-adenovirus vecto
glucocorticoid optimizes its gene transduction
Unpublished
2 (bases 1 to 299)
MURATA,T.
Direct Submission
Submitted (08-DEC-2003) Takahisa Murata, The Universit
Department of Veterinary Pharmacology, Graduate School
Agriculture and Life Sciences; 1-1-1 Yayoi, Bunkyo-ku,
113-8657, Japan (E-mail:murata@mail.v.m.a.u-tokyo.ac.jp
Tel:81-3-5841-5394, Fax:81-3-5841-8183)
Location/Qualifiers
1..299
/organism='Oryctolagus cuniculus'
/mol_type='mRNA'
/db_xref='taxon:9986'
1..299
/gene='lac Z'
<1..>299
/gene='lac Z'
/codon_start=1
/product='beta-D-galactosidase'
/protein_id='BAD04918.1'
/db_xref='GI:39725590'
/translation='VNRHEHPLHGQVMDEQTMVDITLLMKKNFF
WYTLCDRYGLVVDENATEHGMVPMRLTDDPRRLPAMSERVTRF
Query Match 36.9%; Score 28.4; DB 4; Length 299;
Best Local Similarity 96.7%; Pred. No. 7.6;
Matches 29; Conservative 0; Mismatches 1; Indels 0
QY 23 CTAGCATCATCTCTGCATGGTCAGGTGCAT 52
|||||
DB 12 CGAGCATCATCTCTGCATGGTCAGGTGCAT 41
|||||
RESULT 38
ECO487615
LOCUS
DEFINITION
Escherichia coli partial lacZ gene for beta-D-galactos:
mutant 645.
ACCESSION
AJ487615
VERSION
AJ487615.1 GI:21104348
KEYWORDS
beta-D-galactosidase; lacZ gene.
SOURCE
Escherichia coli
ORGANISM
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enteroba
Enterobacteriaceae; Escherichia.
1
Laden,J.C., Philibert,P., Torreilles,F., Pugniere,M. ar
Martineau,P.
Expression and folding of an antibody fragment selecte
high expression levels in Escherichia coli cytoplasm
Unpublished
2 (bases 1 to 1063)
Martineau,P.
Direct Submission

```

mitted (06-MAY-2002) Martineau P., UMR5094 Faculte de Pharmacie,  
s, 15 av. Charles Flahault BP14491, 34093 Montpellier Cedex 5,  
NCE

## Location/Qualifiers

```

1..1063
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
/clone="AMEF mutant 645"
1..1063
/gene="lacZ"
<1..>1063
/gene="lacZ"
/codon_start=2
/evidence=experimental
/transl_table=11
/product="beta-D-galactosidase"
/protein_id="CAD32242.1"
/db_xref="GI:21104349"
/db_xref="GOA:Q8L1L8"
/db_xref="SPTREMBL:Q8L1L8"
/translation="GSYLEDDQMMRMSGIFRDVSLHKKPTQISDFHVATRPNDPFSR
AVLEAEVQMCGLRDYLRVTVSLWQGETQVAGSTAPFGGEIIDERGGYADRVTLNLY
ENPKLWSAEIPNLRYAVVELTADGTLEAEACDVGFREVIENGLILNKGKLLIRG
VNRHHPHPLHQVMDEQTMVQDILLKONNFNAVRCSHYPNHPHPLW
EANTLEHGMVPMNLTDDPRMLPAMSERVTRMVDORNRHPSVILWSLGNESGHGANHD
ALYRWIKSVDPSPRVQVEGGADTTAIDICPMYARVDEDDQFFPAVFNKSIKKWLSLP
GETRPLLCYEAHAMGNSLG"
42
/gene="lacZ"
/note="wild-type"
/replace="g"

```

36.9%; Score 28.4; DB 1; Length 1063;

Similarity 96.7%; Pred. No. 7.8;

Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCTCTGCATGTCAGGTTCAT 52

TAGCATCATCTCTGCATGTCAGGTTCAT 522

1487616 1063 bp DNA linear BCT 21-MAY-2002  
herichia coli partial lacZ gene for beta-D-galactosidase, AMEF  
ant 959.

87616

87616.1 GI:21104350

a-D-galactosidase; lacZ gene.

herichia coli

herichia coli

teria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

erobacteriaceae; Escherichia.

en,J.C., Philibert,P., Torreilles,F., Pugnieri,M. and  
tineau,P.

h expression and folding of an antibody fragment selected in vivo for  
ublished

(bases 1 to 1063)

tineau,P.

ect Submission

mitted (06-MAY-2002) Martineau P., UMR5094 Faculte de Pharmacie,  
s, 15 av. Charles Flahault BP14491, 34093 Montpellier Cedex 5,

NCE

## Location/Qualifiers

```

1..1063
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
/clone="AMEF mutant 959"

```

gene

CDS

```

1..1063
/gene="lacZ"
<1..>1063
/gene="lacZ"
/codon_start=2
/evidence=experimental
/transl_table=11
/product="beta-D-galactosidase"
/protein_id="CAD32243.1"
/db_xref="GI:21104351"
/db_xref="GOA:Q8KHF9"
/db_xref="SPTREMBL:Q8KHF9"
/translation="GSYLEDDQMMRMSGIFRDVSLHKKPTTQISD
AVLEAEVQMCGLRDYLRVTVSLWQGETQVAGSTAPFGGEIIDER
ENPKLWSAEIPNLRYAVVELTADGTLEAEACDVGFREVIENGL
VNRHHPHPLHQVMDEQTMVQDILLKONNFNAVRCSHYPNHPHPLW
EANTLEHGMVPMNLTDDPRMLPAMSERVTRMVDORNRHPSVILW
ALYRWIKSVDPSPRVQVEGGADTTAIDICPMYARVDEDDQFFPA
GETRPLLCYEAHAMGNSLG"
494
/gene="lacZ"
/note="wild-type"
/replace="g"

```

variation

ORIGIN

Query Match 36.9%; Score 28.4; DB 1; Length 1063;  
Best Local Similarity 96.7%; Pred. No. 7.8;  
Matches 29; Conservative 0; Mismatches 1; Indels 0

QY 23 CTAGCATCATCTCTGCATGTCAGGTTCAT 52

Db 493 CAAGCATCATCTCTGCATGTCAGGTTCAT 522

RESULT 40

ECO487617

LOCUS

DEFINITION

Escherichia coli partial lacZ201 gene for beta-D-galact

AMEF mutant W6106.

ACCESSION

AJ487617

VERSION

AJ487617.1 GI:21104352

KEYWORDS

beta-D-galactosidase; lacZ201 gene.

SOURCE

Escherichia coli

ORGANISM

Escherichia coli

REFERENCE

1

AUTHORS

Laden,J.C., Philibert,P., Torreilles,F., Pugnieri,M. ar

Martineau,P.

Expression and folding of an antibody fragment selecte

high expression levels in Escherichia coli cytoplasm

Unpublished

REFERENCE

2 (bases 1 to 1063)

AUTHORS

Martineau,P.

TITLE

Direct Submission

JOURNAL

Submitted (06-MAY-2002) Martineau P., UMR5094 Faculte c

Chrs, 15 av. Charles Flahault BP14491, 34093 Montpellie

FRANCE

FEATURES

source

1..1063

/organism="Escherichia coli"

/mol\_type="genomic DNA"

/db\_xref="taxon:562"

/clone="AMEF mutant W6106"

1..1063

/gene="lacZ201"

<1..>1063

/gene="lacZ201"

/codon\_start=2

/evidence=experimental

/transl\_table=11

/product="beta-D-galactosidase"

/protein\_id="CAD32244.1"

```

/db_xref="GI:21104353"
/db_xref="GOA:Q8KHF9"
/db_xref="SPTREMBL:Q8KHF9"
/translation="GSYLEDDQMMRMGIFRDVSLHKPTTQISDFHVATRFNDDFSR
AVLEAEVQMGELROYLRYVTVSLWQETQVAGTAPFGGEIIDERGGVADRVTLRLNV
ENPKLMSAEIPNLRYAVVELHTADGTLIEAEACDVGFREVIENGLILLNGKPLLIRG
VNRKHHPLHGVMDQETMVQDILLMKQNNFNAVRCSHYPNHPLMYTLCDRYGLYVD
EANIETHGVFNRLTDDPRWLPFMSERTVNVQDRNHPSVVIWSLGNESGHANHD
ALYRWIKSVDPSPRPVQYEGGADTTATDIIICPMYARVDEDDQFFPAVPKWSIKKWLSP
GETRELIICEYAHAMGNSLG"
494
/gene="lacZ201"
/note="wild-type"
/replace="g"

```

```

36.9%; Score 28.4; DB 1; Length 1063;
similarity 96.7%; Pred.No. 7.8;
; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TTAGCATCATCCTCTGCATGTCAGGTCAT 52
|||||
TAAGCATCATCCTCTGCATGTCAGGTCAT 522

```

ed: April 13, 2004, 19:12:11  
9.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

nucleic search, using sw model

April 13, 2004, 18:17:02 ; Search time 1853.5 Seconds  
(without alignments)  
1240.565 Million cell updates/sec

US-10-090-326-24  
77

1 agacaatcacagtctctgcg.....aacccgtggaggacaaggat 77

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

27513289 seqs, 14931090276 residues

hits satisfying chosen parameters: 55026578

length: 0

length: 2000000000

: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

BST.\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_estc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

is the number of results predicted by chance to have a  
alter than or equal to the score of the result being printed,  
rived by analysis of the total score distribution.

# SUMMARIES

Query Match	Length	DB ID	Description
58.4	383	12	BM752895
58.4	412	10	BE696948
56.4	196	10	BF755519
56.4	237	12	BG950112

5	43.4	56.4	284	10	BF818092	BF8180
6	43.4	56.4	295	10	BE697671	BE6976
7	43.4	56.4	302	12	BG957601	BG9576
8	43.4	56.4	401	12	BG981406	BG9814
9	43.4	56.4	417	12	BG955581	BG9555
10	43.4	56.4	524	10	AW362544	AW3625
11	43.4	56.4	540	10	AW937852	AW9378
12	43.4	56.4	556	10	BF759065	BF7590
13	43.4	56.4	556	12	BG956260	BG9562
14	43.4	56.4	563	10	BF805557	BF8055
15	43.4	56.4	579	10	AW604315	AW6043
16	43.4	56.4	593	12	BM742895	BM7428
17	43.4	56.4	596	10	AW858762	AW8587
18	43.4	56.4	601	12	BM786353	BM7863
19	43.4	56.4	612	10	AW360942	AW3609
20	43.4	56.4	644	10	AW853522	AW8535
21	43.4	56.4	904	13	BQ685640	BQ6856
22	43.4	56.4	907	13	BQ689619	BQ6896
23	41.8	54.3	372	10	AW360989	AW3609
24	41.8	54.3	402	10	BF756277	BF7562
25	41.8	54.3	444	12	BM836580	BM8365
26	41.6	54.0	423	9	AA055605	AA0556
27	41	53.2	312	9	AA366575	AA3665
28	41	53.2	389	9	AA132613	AA1326
29	40.2	52.2	287	10	BF817244	BF8172
30	40.2	52.2	327	9	AA134992	AA1349
31	40.2	52.2	630	13	BX480185	BX4801
32	38.4	49.9	494	10	BF084626	BF0846
33	38.4	49.9	623	10	AW862017	AW8620
34	37.6	48.8	440	9	AA127122	AA1271
35	37	48.1	541	10	AW937770	AW9377
36	35.2	45.7	590	10	AW361304	AW3613
37	35	45.5	296	12	BG950148	BG9501
38	34.6	44.9	914	10	BE735665	BE7356
39	34.4	44.7	237	9	AA327327	AA3273
40	34.4	44.7	423	10	AW849404	AW8494
41	33.6	43.6	338	10	AW383654	AW3836
42	33.2	43.1	577	13	BX644443	BX6444
43	32.8	42.6	233	10	AW383495	AW3834
44	32.8	42.6	301	12	BG950114	BG9501
45	32.8	42.6	321	10	BF336900	BF3369

## ALIGNMENTS

RESULT 1  
BM752895  
LOCUS  
DEFINITION K-EST0029613 S9SNU601 Homo sapiens cdna clone S9SNU601-  
383 bp mRNA linear EST  
K-EST0029613 S9SNU601 Homo sapiens cdna clone S9SNU601-  
mRNA sequence.  
ACCESSION BM752895.1 GI:19082513  
VERSION  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho  
1 (bases 1 to 383)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., K  
Kim,Y.S.  
TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 8 row: G column: 11  
High quality sequence stop: 383.

Location/Qualifiers  
 1. 383  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S9SNU601-8-G11"  
 /sex="M"  
 /tissue\_type="Ascites"  
 /cell\_type="Epithelial"  
 /cell\_line="SNU-601"  
 /lab\_host="Top10F"  
 /clone\_lib="S9SNU601"  
 /note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI; Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including SfiI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized with Superscript II using SfiI oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into DraIII- digested pME18S-FL3 vector. The obtained cDNA vectors were used for transfection of competent cells E. coli Top10F, by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

Query Match 58.4%; Score 45; DB 12; Length 383;  
 Best Local Similarity 74.0%; Pred. No. 0.00093;  
 Matches 0; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGGGAGGATCATCTCTGCATGTCAGTGCATATCTCCAAAC 60  
 AGACAATCACAGTCTCTGGGAGGATGAGCCAGCCCTCCATCTCCAGCAAAACTCCAAAC 337  
 TCGTGGAGGACAGGAT 77  
 CCGTGGAGGACAGGAT 354

596948 412 bp mRNA linear EST 11-SEP-2000  
 1-CT0408-220700-011-f06 CT0408 Homo sapiens cDNA, mRNA sequence.  
 596948  
 596948.1 GI:10084108  
 F.  
 no sapiens (human)  
 no sapiens  
 karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 (bases 1 to 412)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Bordon, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.E., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J., Simpson, A.J.G.  
 Shotgun sequencing of the human transcriptome with ORF  
 puce tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 202663  
 737800  
 atact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human C  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&tl=311000-002-f10&t3=2000-10-31&t4=1)  
 Seg primer: puc 18 forward  
 High quality sequence stop: 196.

This sequence was derived from the FAPESP/LICR Human C  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&tl=311000-002-f10&t3=2000-07-22&t4=1)  
 Seg primer: puc 18 forward  
 High quality sequence start: 8  
 High quality sequence stop: 412.  
 Location/Qualifiers  
 1. 412  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="CT0408"  
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; A mini-library was made by cloning prod from ORESSES PCR (U.S. Letters Patent applic 196,716 - Ludwig Institute for Cancer Research into the pUC 18 vector. Reverse transcriptio mRNA and cDNA amplification were performed un stringency conditions."

Query Match 58.4%; Score 45; DB 10; Length 412;  
 Best Local Similarity 74.0%; Pred. No. 0.00096;  
 Matches 57; Conservative 0; Mismatches 20; Indels 0

QY 1 AGACAATCACAGTCTCTGGGAGGATCATCTCTGCATGTCAGTGCATATCTCCAAAC  
 DB 331 AGACAATCACAGTCTCTGGGAGGATGAGCCAGCCCTCCATCTCCAGCAACAA  
 QY 61 CCGTGGAGGACAGGAT 77  
 DB 271 CCGTGGAGGACAGGAT 255

RESULT 3  
 BF755519/c  
 LOCUS BF755519 196 bp mRNA linear ES  
 DEFINITION MR2-CT0595-311000-002-f10 CT0595 Homo sapiens cDNA, mRNA  
 ACCESSION BF755519  
 VERSION BF755519.1 GI:12103419  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 (bases 1 to 196)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Bordon, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.E., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J., Simpson, A.J.G.  
 Shotgun sequencing of the human transcriptome with ORF  
 puce tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 10737800  
 PUBMED 10737800  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human C  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&tl=311000-002-f10&t3=2000-10-31&t4=1)  
 Seg primer: puc 18 forward  
 High quality sequence stop: 196.

```

Location/Qualifiers
1..196
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0595"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

56.4%; Score 43.4; DB 10; Length 196;
Similarity 72.7%; Pred. No. 0.0022;
; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

AGACATCACAGTCTCTGGGAGCATCATCTCTGCGATGGTCAGTCAACTCAAAC 60
AGACATCACAGTCTCTGGGAGCATCATCTCTGCGATGGTCAGTCAACTCAAAC 116

>CGTGGAGGACAAAGGAT 77
|||||
>CGTGGAGGACAAAGGAT 99
|||||

p50112 237 bp mRNA linear EST 12-JUN-2001
!-CT0639-030101-691-a04 CT0639 Homo sapiens cDNA, mRNA sequence.
p50112
p50112.1 GI:14368283
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 237)
s Neto, E., Garcia Corrae, R., Verjovski-Almeida, S., Briones, M.R.,
ai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
dman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
nstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
are, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
pson, A.J.
tgun sequencing of the human transcriptome with ORF expressed
uence tags
c. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
02663
37800
tact: Simpson A.J.G.
ratory of Cancer Genetics
wig Institute for Cancer Research
Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
zil
: +55-11-2704922
il: asimpson@ludwig.org.br
s sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
tp://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM2&t2=CM2-CT0639-
101-691-a04&t3=2001-01-03&t4=1)
primer: puc 18 forward
h quality sequence start: 12
h quality sequence stop: 237.
Location/Qualifiers
1..237
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

```

```

/clone_lib="CT0639"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

Query Match 56.4%; Score 43.4; DB 12; Length 237;
Best Local Similarity 72.7%; Pred. No. 0.0024;
Matches 56; Conservative 0; Mismatches 21; Indels 0

QY 1 AGACATCACAGTCTCTGGGAGCATCATCTCTGCGATGGTCAGTCAACTCAAAC
|||||
Db 93 AGACATCACAGTCTCTGGGAGCATCATCTCTGCGATGGTCAGTCAACTCAAAC
|||||
QY 61 CCGTGGAGGACAAAGGAT 77
|||||
Db 153 CCGTGGAGGACAAAGGAT 169
|||||

RESULT 5
BF818092 284 bp mRNA linear EST
LOCUS MR3-CI0184-151200-007-c04 CI0184 Homo sapiens cDNA, mRNA
DEFINITION BF818092
ACCESSION BF818092
VERSION BF818092.1 GI:12155181
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Eumammalia; Eutheria; Primates; Catarrhini; Hominidae; Hominidae; Homo.
(bases 1 to 284)
Dias Neto, E., Garcia Corrae, R., Verjovski-Almeida, S., Briones, M.R.,
ai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
dman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
nstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
are, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
pson, A.J.
Shotgun sequencing of the human transcriptome with ORF
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR3&t2=
151200-007-c04&t3=2000-12-15&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 284.
Location/Qualifiers
1..284
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CI0184"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

```

## FEATURES

source



DEFINITION CM4-CT06556-130201-852-c02 CT06556 Homo sapiens cDNA, mRNA  
ACCESSION BG955581  
VERSION BG955581.1 GI:14373752  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo (bases 1 to 417)  
 Dias Neco, E., Garcia Correa, R., Verjovski-Almeida, S., B. Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., S. Brunstein, A., deOliveira, P.S., Bucher, P., Jorgensen, C.V. O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F.F., de Souza, J. Simpson, A.J.

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., S. Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V. O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, A.J., Simpson, A.J.

Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza  
Simpson, A.J.

TITLE	Shotgun sequencing of the human transcriptome with ORF sequence tags
JOURNAL	PROC. NATL. ACAD. SCI. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
PUBMED	10737800
COMMENT	Contact: Simpson A. J. G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Brazil

Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao  
Brazil

**Tel: +55-11-2704922**  
**Fax: +55-11-2707001**

Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Ca  
Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.p?71t1=CM4&t2=1302011-652-C02&t3=2001-02-13&t4=1>)  
Seq primer: puc 18 forward  
High quality sequence start: 34  
High quality sequence stop: 413.

```
Seq primer: puc 18 forward
High quality sequence start: 34
High quality sequence stop: 413.
Location/Qualifiers
1..417
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev stage="Adult"
```

```

/clone_lib="CR0656"
/notes="Organ: colon; Vector: puc18; Site 1: Sm
SmaI; A mini-library was made by cloning produ
from ORESTES PCR (U.S. Letters Patent applica
196,716 - Ludwig Institute for Cancer Research
into the pUC 18 vector. Reverse transcription
mRNA and cDNA amplification were performed und
stringency conditions."

```

QY 1 AGCAATCACAGTCTCTGGGAAGCATCATCTCTGTCATGTCAGGTCAATAA  
Db 367 AGCAATCACAGTCTCTGGGAGTGCCCAAGCCCTCCATCTCCAGCAACAAC  
QY 61 CCGTGGAGGACAAGGAT 77

Db 307 CCGTGGAGCACAGGAT 291

RESULT 10	524 bp	linear	EST
AW362544/c	AW362544	mrna	
LOCUS			
DEFINITION	RC3-CT0283-201199-011-g06	CT0283	Hom sapiens cDNA, mRNA

ACCESSION AW3622544  
VERSION AW362544..1 GI:6867194  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. (bases 1 to 524)  
 P http://www.ludwig.org.br/ORESTES.  
 FAPESP/LICR Human Cancer Genome Project  
 ublished (1999)  
 tact: Simpson A.J.G.  
 oratory of Cancer Genetics  
 wig Institute for Cancer Research  
 Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, zil  
 : +55-11-2704922  
 : +55-11-2707001  
 il: asimpson@ludwig.org.br  
 s sequence was derived from the FAPESP/LICR Human Cancer Genome ject. This entry can be seen in the following URL  
 tp://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3&t2=RC3-CT0283-199-011-g06&t3=1999-11-20&t4=1)  
 199-011-g06&t3=1999-11-20&t4=1)  
 h quality sequence start: 12  
 h quality sequence stop: 524.  
 Location/Qualifiers  
 1..524  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="CT0283"  
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
 milarity 56.4%; Score 43.4; DB 10; Length 524;  
 milarity 72.7%; Pred. No. 0.0034;  
 Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
 GACAAATCACAGTCTCGCGAAGCATCATCTCTGCATGTCAGGTCATCAACTCCAAAC 60  
 GACAAATCACAGTCTCGCGAGTGCCTCCAGCCCTCCATCTCCAGCAACATCCAAAC 202  
 :CGTGGAGGACAAAGGAT 77  
 :CGTGGAGGACAAAGGAT 185  
 :CGTGGAGGACAAAGGAT 185  
 137852 540 bp mRNA linear EST 30-MAY-2000  
 i-DT0045-010300-108-f02 DT0045 Homo sapiens cDNA, mRNA sequence.  
 137852  
 137852.1 GI:8113281  
 F.  
 no sapiens (human)  
 no sapiens  
 raryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. (bases 1 to 540)  
 as Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., ai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., ldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., instein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and pson,A.J.  
 ctgun sequencing of the human transcriptome with ORF expressed uence tags  
 oc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 202663  
 202663  
 737800

## COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Ca project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=QV300-1C8-f02&t3=2000-03-01&t4=1)  
 300-1C8-f02&t3=2000-03-01&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 9  
 High quality sequence stop: 540.  
 Location/Qualifiers

## FEATURES

source  
 1..540  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="DT0045"  
 /note="Organ: denis drash; Vector: puc18; Site Site\_2: SmaI; A mini-library was made by cloni derived from ORESTES PCR (U.S. Letters Patent No. 196,716 - Ludwig Institute for Cancer Res profiles into the pUC 18 vector. Reverse tran tissue mRNA and cDNA amplification were perfor low stringency conditions."

## ORIGIN

Query Match 56.4%; Score 43.4; DB 10; Length 540;  
 Best Local Similarity 72.7%; Pred. No. 0.0034;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0;  
 QY 1 AGACAATCACAGTCTCGCGAAGCATCATCTCTGCATGTCAGGTCATCAACTCCAAAC  
 Db 378 AGACAATCACAGTCTCTCGCGAGTGCCTCCAGCCCTCCATCTCCAGCAACAAAC  
 QY 61 CCGTGGAGGACAAAGGAT 77  
 Db 438 CCGTGGAGGACAAAGGAT 454

## RESULT 12

BF759065 556 bp mRNA linear EST  
 LOCUS  
 DEFINITION QV4-CT0069-281100-604-h04 CT0069 Homo sapiens cDNA, mRNA  
 ACCESSION BF759065  
 VERSION BF759065.1 GI:12106965  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hc  
 1 (bases 1 to 556)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., F Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Cost Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., s Bruns-ein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de S Simpson,A.J.

## REFERENCE

Shotgun sequencing of the human transcriptome with ORF sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE  
 20202563  
 10737300  
 PUBMED

## TITLE

Shotgun sequencing of the human transcriptome with ORF sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE  
 20202563  
 10737300  
 PUBMED

## COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Brazil  
 Tel: +55-11-2704922

```

i: +55-11-2707001
il: asimpson@ludwig.org.br
s sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
tp://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV4&t2=QV4-CT0069-
100-604-b04&t3=2000-11-28&t4=1)
primer: puc 18 forward
n quality sequence start: 223
n quality sequence stop: 555.
Location/Qualifiers
1. .556
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0069"
/notes="Organ: colon; Vector: pUC18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
56.4%; Score 43.4; DB 10; Length 556;
milarity 72.7%; Pred.No. 0.0035; Indels 0; Gaps 0;
Conservative 0; Mismatches 21;
GACAAATCACAGTCTCTGGGAGGACATCATCTCTGCATGGTCAGTCAATCAATCCAAAC 60
|||||
GACAAATCACAGTCTCTGGGAGGTCGCCAAGCCCTCCATCTCCAGCAACAATCCAAAC 422
|||||
CTGGAGGACAAGGAT 77
|||||
CTGGAGGACAAGGAT 439
|||||

56260 56260.1 GI:14374431
-CT0616-190201-553-b02 CT0616 Homo sapiens cdNA, mRNA sequence.
56260
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 556)
s Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
ai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
dman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
nstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
are, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
pson, A.J.
tgun sequencing of the human transcriptome with ORF expressed
uence tags
c. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
02663
37800
tact: Simpson A.J.G.
oratory of Cancer Genetics
wig Institute for Cancer Research
Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
zil
: +55-11-2704922
: +55-11-2707001
il: asimpson@ludwig.org.br
s sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
tp://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&t2=QV3-CT0616-
201-553-b02&t3=2001-02-19&t4=1)

```

```

Seq primer: puc 18 forward
High quality sequence stop: 549.
Location/Qualifiers
1. .556
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0616"
/notes="Organ: colon; Vector: pUC18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
56.4%; Score 43.4; DB 12; Length 556;
Best Local Similarity 72.7%; Pred.No. 0.0035; Indels 0;
Matches 56; Conservative 0; Mismatches 21;
ORIGIN
Query Match 56.4%; Score 43.4; DB 12; Length 556;
Best Local Similarity 72.7%; Pred.No. 0.0035; Indels 0;
Matches 56; Conservative 0; Mismatches 21;
QY 1 AGACAATCACAGTCTCTGGGAGGACATCATCTCTGCATGGTCAGTCAATCAATCAAC
|||||
Db 284 AGACAATCACAGTCTCTGGGAGGTCGCCAAGCCCTCCATCTCCAGCAACAAC
|||||
QY 61 CCGTGAGGACAAGGAT 77
|||||
Db 344 CCGTGAGGACAAGGAT 360
|||||

RESULT 14
BF05557/c
LOCUS
DEFINITION
ACCESSION BF05557
VERSION BF05557.1 GI:12134546
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
1 (bases 1 to 563)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., B
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Cost
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., S
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de So
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF
Sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Ca
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV1&t2
06110C-462-g02&t3=2000-11-06&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 563.
Location/Qualifiers
1. .563
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```

```

/dev_stage="Adult"
/clone_lib="C10173"
/note="Organ: colon ins; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

```

56.4%; Score 43.4; DB 10; Length 563;
milarity 72.7%; Pred. No. 0.0035;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;
GACATCACAGTCTCTCGGAGCAGCATCTCTGTCATGTCAGGTCATCAATCCAAAC 60
|||||
GACATCACAGTCTCTCGGAGCAGCCTCCATCTCCAGCACAATCCAAAC 162
|||||
CGTGGAGGACAAGGAT 77
|||||
CGTGGAGGACAAGGAT 145
|||||

```

```

04315 579 bp mRNA linear EST 23-MAR-2000
-CT0219-280100-062-G01 CT0219 Homo sapiens cDNA, mRNA sequence.
04315
04315.1 GI:7309056
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(Bases 1 to 579)
P http://www.ludwig.org.br/ORESTES.
P FAPESP/LICR Human Cancer Genome Project
ublished (1999)
tact: Simpson A.J.G.
ortory of Cancer Genetics
wlg Institute for Cancer Research
Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
uzil
: +55-11-2704922
: +55-11-2707001
il: asimpson@ludwig.org.br
s sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
tp://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&t2=IL3-CT0219-
100-062-G01&t3=2000-01-28&t4=1)
primer: puc 18 forward
h quality sequence stop: 572.
Location/Qualifiers
1..579
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0219"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
56.4%; Score 43.4; DB 10; Length 579;
milarity 72.7%; Pred. No. 0.0035;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```

```

QY 1 AGACAATCACAGTCTCTCGGAGCAGCATCTCTGTCATGTCAGGTCATCAATCAAC
|||||
DB 471 AGACAATCACAGTCTCTCGGAGCAGCCTCCATCTCCAGCACAAC
|||||
QY 61 CGTGGAGGACAAGGAT 77
|||||
DB 531 CGTGGAGGACAAGGAT 547
|||||

```

```

RESULT 16
BM742895 593 bp mRNA linear EST
K-EST0015932 S6SNU620 Homo sapiens cDNA clone S6SNU620-
RNA sequence.
BM742895
BM742895.1 GI:19064224
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
1 (bases 1 to 593)
Kim,N.S., Rahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., K
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 10 row: G column: 05
High quality sequence stop: 593.
Location/Qualifiers
1..593
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S6SNU620-10-G05"
/sex="F"
/tissue_type="Ascites"
/cell_type="Scattering floating"
/lab_host="SNU-620"
/clone_lib="S6SNU620"
/note="Organ: Stomach; Vector: pCNS; Site 1: I
Site 2: NotI; The poly (A) + RNA was dephospho
bacterial alkaline phosphatase (BAP) and then
with tabacco acid pyrophosphatase (TAP). The
intact mRNA was ligated with DNA-RNA linker in
I site by treatment of T4 RNA ligase and the
cDNA was synthesized from oligo dt-selected m
priming with dt-tailed vector. The dt-tailed
adjusted to have about 60nt. The cDNA vector
circularized with E. coli DNA ligase after dig
EcoRI which site is also included in vector.
converted to a DNA strand by Okayama-Berg met
obtained cDNA vectors were used for transfor
competent cells E. coli Top10P by electropor
The cDNA libraries constructed by this method
full-length enriched cDNA library."

```

```

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 10 row: G column: 05
High quality sequence stop: 593.
Location/Qualifiers
1..593
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S6SNU620-10-G05"
/sex="F"
/tissue_type="Ascites"
/cell_type="Scattering floating"
/lab_host="SNU-620"
/clone_lib="S6SNU620"
/note="Organ: Stomach; Vector: pCNS; Site 1: I
Site 2: NotI; The poly (A) + RNA was dephospho
bacterial alkaline phosphatase (BAP) and then
with tabacco acid pyrophosphatase (TAP). The
intact mRNA was ligated with DNA-RNA linker in
I site by treatment of T4 RNA ligase and the
cDNA was synthesized from oligo dt-selected m
priming with dt-tailed vector. The dt-tailed
adjusted to have about 60nt. The cDNA vector
circularized with E. coli DNA ligase after dig
EcoRI which site is also included in vector.
converted to a DNA strand by Okayama-Berg met
obtained cDNA vectors were used for transfor
competent cells E. coli Top10P by electropor
The cDNA libraries constructed by this method
full-length enriched cDNA library."

```

```

FEATURES
source
1..593
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S6SNU620-10-G05"
/sex="F"
/tissue_type="Ascites"
/cell_type="Scattering floating"
/lab_host="SNU-620"
/clone_lib="S6SNU620"
/note="Organ: Stomach; Vector: pCNS; Site 1: I
Site 2: NotI; The poly (A) + RNA was dephospho
bacterial alkaline phosphatase (BAP) and then
with tabacco acid pyrophosphatase (TAP). The
intact mRNA was ligated with DNA-RNA linker in
I site by treatment of T4 RNA ligase and the
cDNA was synthesized from oligo dt-selected m
priming with dt-tailed vector. The dt-tailed
adjusted to have about 60nt. The cDNA vector
circularized with E. coli DNA ligase after dig
EcoRI which site is also included in vector.
converted to a DNA strand by Okayama-Berg met
obtained cDNA vectors were used for transfor
competent cells E. coli Top10P by electropor
The cDNA libraries constructed by this method
full-length enriched cDNA library."

```

```

ORIGIN
Query Match 56.4%; Score 43.4; DB 12; Length 593;
Best Local Similarity 72.7%; Pred. No. 0.0036;
Matches 56; Conservative 0; Mismatches 21; Indels 0
1 AGACAATCACAGTCTCTCGGAGCAGCATCTCTGTCATGTCAGGTCATCAAC
|||||
240 AGACAATCACAGTCTCTCGGAGCAGCCTCCATCTCCAGCACAAC
|||||

```

```

QY 1 AGACAATCACAGTCTCTCGGAGCAGCATCTCTGTCATGTCAGGTCATCAAC
|||||
DB 240 AGACAATCACAGTCTCTCGGAGCAGCCTCCATCTCCAGCACAAC
|||||

```



GTGGAGGACAAGGAT 77  
|||||  
GTGGAGGACAAGGAT 316

8762 596 bp mRNA linear EST 19-MAY-2000  
CT0347-160200-013-h10 CT0347 Homo sapiens cDNA, mRNA sequence.

8762 1 GI:7954455

: sapiens (human)

: sapiens  
ryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
alia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

bases 1 to 596)

Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
i,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
man,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
stein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,  
re,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
son,A.J.

gun sequencing of the human transcriptome with ORF expressed

ence tags

: Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

12663

7800

act: Simpson A.J.G.

ratory of Cancer Genetics

ig Institute for Cancer Research

Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

il

+55-11-2704922

+55-11-2707001

.l: asimpson@ludwig.org.br

: sequence was derived from the FAPESP/LICR Human Cancer Genome  
lect. This entry can be seen in the following URL  
p://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=RC3-CT0347-160  
013-h10&t3=2000-02-16&t4=1

primer: puc 18 forward

quality sequence stop: 593.

Location/Qualifiers

1..596

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="CT0347"

/note="Organ: colon; Vector: puc18; Site\_1: SmaI; Site\_2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No.  
196,716 - Ludwig Institute for Cancer Research) profiles  
into the puc 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

56.4%; Score 43.4; DB 10; Length 596;

ilarity 72.7%; Pred. No. 0.0036;

Conservative 0; Mismatches 21; Indels 0; Gaps 0;

JACAATCACAGTCTCTCGGAGAGCATCATCTCTGCATGGTCAGGTCATAATCCAAAC 60

|||||

JACAATCACAGTCTCTCGGAGAGTCCCAAGCCCTCCATCTCCAGCAACAATCCAAAC 154

|||||

CGTGGAGGACAAGGAT 77

|||||

CGTGGAGGACAAGGAT 171

BM786353

LOCUS

DEFINITION

K-EST0065059 S9SNU601 Homo sapiens

mRNA sequence.

ACCESSION

BM786353

VERSION

BM786353.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo

1 (bases 1 to 601)

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y.

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,

Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 36 row: E column: 10

High quality sequence stop: 601.

Location/Qualifiers

1..601

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S9SNU601-36-E10"

/sex="M"

/tissue\_type="Ascites"

/cell\_type="Epithelial"

/cell\_line="SNU-601"

/lab\_host="Top10P"

/clone\_lib="S9SNU601"

/note="Organ: Stomach; Vector: pME18-FL3; Site:

Site 2: XhoI; The poly (A)+ RNA was dephosphor-

bacterial alkaline phosphatase (BAP) and then

with tabacco acid pyrophosphatase (TAP). The d

intact mRNA was ligated with DNA-RNA linker in

site by treatment of T4 RNA ligase and the fir

cDNA was synthesized with Superscript II using

oligo-dT primer. After first strand synthesis,

degraded by NaOH treatment and cDNA was amplif

reaction. The PCR products were digested with

cloned into DraIII- digested pME18S-FL3 vector

obtained cDNA vectors were used for transforma

competent cells E. coli Top10P, by electropora

The cDNA libraries constructed by this method

full-length enriched cDNA library."

ORIGIN

Query Match

Best Local Similarity

Matches 56; Conservative 0; Mismatches 21; Indels 0;

56.4%; Score 43.4; DB 12; Length 601;

1 AGACAATCACAGTCTCTCGGAGAGCATCATCTCTGCATGGTCAGGTCATAAC

|||||

500 AGACAATCACAGTCTCTCGGAGAGTCCCAAGCCCTCCATCTCCAGCAACAAC

|||||

61 CCGTGAGGACAAGGAT 77

|||||

560 CCGTGAGGACAAGGAT 576

|||||

RESULT 19

AW360942/c

LOCUS

PM1-CT0247-151299-005-e09 CT0247 Homo sapiens

DEFINITION

AW360942

Accession

612 bp mRNA linear EST

PM1-CT0247-151299-005-e09 CT0247 Homo sapiens cDNA, mRN

Accession

AW360942

60942.1 GI:6865592  
 o sapiens (human)  
 arvota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 (bases 1 to 612)  
 P http://www.ludwig.org.br/ORESTES.  
 FAPESP/LICR Human Cancer Genome Project  
 ublished (1999)  
 tact: Simpson A.J.G.  
 oratory of Cancer Genetics  
 wig Institute for Cancer Research  
 Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 zil  
 : +55-11-2704922  
 : +55-11-2707001  
 il: asimpson@ludwig.org.br  
 s sequence was derived from the FAPESP/LICR Human Cancer Genome  
 ject. This entry can be seen in the following URL  
 tp://www.ludwig.org.br/scripts/gethtml2.pl?tl=PMI&st2=PMI-CT0247-  
 239-005-e09&t3=1999-12-15&t4=1  
 primer: puc 18 forward  
 h quality sequence start: 24  
 n quality sequence stop: 598.  
 Location/Qualifiers  
 1. 612  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="CT0247"  
 /note="Organ: colon; Vector: puc18; Site 1: Smal; Site 2:  
 Smal; A mini-library was made by cloning products derived  
 from ORESTES PCR (U.S. Letters Patent application No.  
 196,716 - Ludwig Institute for Cancer Research) profiles  
 into the pUC 18 vector. Reverse transcription of tissue  
 mRNA and cDNA amplification were performed under low  
 stringency conditions."  
 56.4%; Score 43.4; DB 10; Length 612;  
 milarity 72.7%; Pred. No. 0.0036;  
 Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
 GACAATCACAGTCTCTCGGAAGCATCATCTCTGCATGTCAGGTGATACCTCCAAAC 60  
 GACAATCACAGTCTCTCGGAGCTGCGCAAGCCCTCCATCTCCAGCAACAATCCAAAC 171  
 :CGTGGAGGACAGGAT 77  
 :CGTGGAGGACAGGAT 154  
 153522 644 bp mRNA linear EST 19-MAY-2000  
 -CT0252-140300-026-906 CT0252 Homo sapiens cDNA, mRNA sequence.  
 153522  
 153522.1 GI:7949267  
 no sapiens (human)  
 no sapiens  
 karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 (bases 1 to 644)  
 as Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 gai,W.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Idman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
 instein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,  
 Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 npson,A.J.  
 ctgum sequencing of the human transcriptome with ORF expressed

sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 MEDLINE  
 10737800  
 PUBMED  
 COMMENT  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sac  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Ca  
 project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=RC  
 300-026-g06&t3=2000-03-14&t4=1)  
 seq primer: puc 18 forward  
 High quality sequence start: 24  
 High quality sequence stop: 615.  
 FEATURES  
 Location/Qualifiers  
 1. 644  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="CT0252"  
 /note="Organ: colon; Vector: puc18; Site 1: Sm  
 Smal; A mini-library was made by cloning produ  
 from ORESTES PCR (U.S. Letters Patent applica  
 196,716 - Ludwig Institute for Cancer Researc  
 into the pUC 18 vector. Reverse transcrip  
 mRNA and cDNA amplification were performed unc  
 stringency conditions."  
 56.4%; Score 43.4; DB 10; Length 644;  
 Query Match 72.7%; Pred. No. 0.0037;  
 Best Local Similarity 72.7%; Mismatches 21; Indels 0;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0;  
 QY 1 AGACAATCACAGTCTCTCGGAAGCATCATCTCTGCATGTCAGGTGATAC  
 |||||  
 Db 360 AGACAATCACAGTCTCTCGGAGCTGCGCAAGCCCTCCATCTCCAGCAACA  
 |||||  
 QY 61 CCGTGGAGGACAGGAT 77  
 |||||  
 Db 300 CCGTGGAGGACAGGAT 284  
 |||||  
 RESULT 21  
 BQ685640  
 LOCUS  
 DEFINITION AGENCOURT\_8345228 NIH\_MGC\_110 Homo sapiens cDNA clone  
 5', mRNA sequence.  
 ACCESSION BQ685640  
 VERSION BQ685640.1 GI:21810956  
 KEYWORDS EST  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H  
 1 (bases 1 to 904)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collectic  
 Unpublished (1999)  
 CONTACT Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LI  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov

ate: L1CM2392 row: b column: 19  
 3h quality sequence stop: 686.  
 Location/Qualifiers  
 1..904  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6249906"  
 /tissue\_type="ductal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_110"  
 /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;  
 Site 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

56.4%; Score 43.4; DB 13; Length 904;  
 imilarity 72.7%; Pred. No. 0.0043;  
 ; Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
 AGACAATCACAGTCTCTCGGAGCATCATCTCTGCATGTCAGTTCATTAATCCAAAC 60  
 AGACATCACAGTCTCTCGGAGCTGCCAAGCCCTCCATCTCCAGCAACTCCAAAC 456  
 CGTGGAGGACAAAGGAT 77  
 |||||||||  
 CGTGGAGGACAAAGGAT 473  
 |||||||||

89619 907 bp mRNA linear EST 15-JUL-2002  
 ENCOURT\_8186629 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6251590  
 mRNA sequence.  
 89619 1 GI:21814935  
 ..  
 to sapiens (human)  
 to sapiens  
 aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 (bases 1 to 907)  
 i-MGC http://mgc.nci.nih.gov/.  
 ional Institutes of Health, Mammalian Gene Collection (MGC)  
 ublished (1999)  
 ctact: Robert Strausberg, Ph.D.  
 il: cgapbs-remail.nih.gov  
 use Procurement: ATCC  
 NA Library Preparation: Rubin Laboratory  
 NA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 A Sequencing by: Agencourt Bioscience Corporation  
 one distribution: MGC clone distribution information can be  
 nd through the I.M.A.G.E. Consortium/LLNL at:  
 p://image.llnl.gov  
 ate: L1CM2396 row: h column: 23  
 3h quality sequence stop: 514.  
 Location/Qualifiers  
 1..907  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6251590"  
 /tissue\_type="ductal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_110"  
 /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;  
 Site 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library co  
 Ling Hong in the laboratory of Gerald M. Rubi  
 of California, Berkeley) using ZAP-cDNA synt  
 (Stratagene) and Superscript II RT (Life Tech  
 Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 56.4%; Score 43.4; DB 13; Length 907;  
 Best Local Similarity 72.7%; Pred. No. 0.0043;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0

QY 1 AGACAATCACAGTCTCTCGGAGCATCATCTCTGCATGTCAGTTCAGTTCATAA  
 |||||||||  
 Db 397 AGACAATCACAGTCTCTCGGAGCTGCCAAGCCCTCCATCTCCAGCAACAA  
 |||||||||  
 QY 61 CCGTGGAGGACAAAGGAT 77  
 |||||||||  
 Db 457 CCGTGGAGGACAAAGGAT 473  
 |||||||||

## RESULT 23

AW360989/c  
 LOCUS AW360989 372 bp mRNA linear ES  
 DEFINITION PM1-CT0247-291299-007-b09 CT0247 Homo sapiens cDNA, mEJ  
 ACCESSION AW360989  
 VERSION AW360989.1 GI:6865639  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H

## REFERENCE

AUTHORS 1 (bases 1 to 372)  
 TITLE HCGP http://www.ludwig.org.br/ORESTES.  
 JOURNAL The FAPESP/LICR Human Cancer Genome Project  
 COMMENT Unpublished (1999)  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01508-010, Sa  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br

## FEATURES

source  
 1..372  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="CT0247"  
 /note="Organ: colon; Vector: puc18; Site 1: S  
 SmaI; A mini-library was made by cloning prod  
 from ORESTES PCR (U.S. Letters Patent applica  
 196,716 - Ludwig Institute for Cancer Researc  
 into the pUC 18 vector. Reverse transcriptio  
 mRNA and cDNA amplification were performed unc  
 stringency conditions."

## FEATURES

source  
 1..372  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="CT0247"  
 /note="Organ: colon; Vector: puc18; Site 1: S  
 SmaI; A mini-library was made by cloning prod  
 from ORESTES PCR (U.S. Letters Patent applica  
 196,716 - Ludwig Institute for Cancer Researc  
 into the pUC 18 vector. Reverse transcriptio  
 mRNA and cDNA amplification were performed unc  
 stringency conditions."

## ORIGIN

Query Match 54.3%; Score 41.8; DB 10; Length 372;  
 Best Local Similarity 71.4%; Pred. No. 0.0094;  
 Matches 55; Conservative 0; Mismatches 22; Indels 0;

QY 1 AGACAATCACAGTCTCTCGGAGCATCATCTCTGCATGTCAGTTCATAA  
 |||||||||  
 Db 212 AGACAATCACAGTCTCTCGGAGCTGCCAAGCCCTCCATCTCCAGCAACAA  
 |||||||||

CGCTGGAGGACAAAGGAT 77  
|||||  
CGCTGGAGGACAAAGGAT 136

756277 402 bp mRNA linear EST 12-JAN-2001  
3-CT0558-301000-453-b11 CT0558 Homo sapiens cDNA, mRNA sequence.

756277  
756277.1 GI:12104177

T.  
no sapiens (human)

no sapiens  
karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

(bases 1 to 402)  
as Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Gai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Idman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
Unstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
Hare, M.J., Soares, F., Brentani, R.R., Reis, I.F., de Souza, S.J. and  
mpson, A.J.

Organ sequencing of the human transcriptome with ORF expressed

quence tags  
cc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

202663  
737800

ntact: Simpson A.J.G.  
boratory of Cancer Genetics

dwig Institute for Cancer Research

a Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
azil

1: +55-11-2704922  
x: +55-11-2707001

all: asimpson@ludwig.org.br

is sequence was derived from the FAPESP/LICR Human Cancer Genome  
ject. This entry can be seen in the following URL:  
http://www.ludwig.org.br/scripts/gethtml2.pl?ti=QV3&t2=QV3-CT0558-  
1000-453-b11&t3=2000-10-30&t4=1

q primer: puc 18 forward  
gh quality sequence stop: 402.

Location/Qualifiers  
1. .402

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev stage="Adult"

/clone\_lib="CT0558"

/notes="Organ: colon; Vector: puc18; Site\_1: SmaI; Site\_2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No.  
196,716 - Ludwig Institute for Cancer Research) profiles  
into the puc 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

54.3%; Score 41.8; DB 10; Length 402;

imilarity 71.4%; Pred. No. 0.0097;

; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

AGCAATCAGTCTCTGGGAGCATATCCCTCTGCGATGGTCAGGTCAATCAAC 60

AGCAATCAGGCTCTGGGAGTGCACCAAGCCCTCCATCTCCAGCAACATCCAAAC 155

CGCTGGAGGACAAAGGAT 77  
|||||  
CGCTGGAGGACAAAGGAT 172

BM836580

LOCUS  
DEFINITION

ACCESSION  
VERSION

KEYWORDS  
SOURCE

ORGANISM

REFERENCE  
AUTHORS

TITLE  
JOURNAL

COMMENT

FEATURES  
source

BM836580 444 bp mRNA linear ES  
K-EST0112341 S9SNU601 Homo sapiens cDNA clone S9SNU601  
mRNA sequence.

BM836580  
BM836580.1 GI:19192989

EST.  
Homo sapiens (human)

Homo sapiens

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H  
1 (bases 1 to 444)

Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.  
Ch, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S.,  
Kim, Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 86 row: H column: 09

High quality sequence stop: 444.

Location/Qualifiers  
1. .444

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S9SNU601-86-H09"

/sex="M"

/tissue\_type="Ascites"

/cell\_type="Epithelial"

/lab\_line="SNU-601"

/lab\_host="Top10F"

/clone\_lib="S9SNU601"

/note="Organ: Stomach; Vector: pME18-FL3; Site  
Site\_2: XhoI; The poly (A)+ RNA was dephospho  
bacterial alkaline phosphatase (BAP) and then  
with tabacco acid pyrophosphatase (TAP). The  
intact mRNA was ligated with DNA-RNA linker i  
site by treatment of T4 RNA ligase and the fi  
cDNA was synthesized with Superscript II usin  
oligo-dt primer. After first strand synthesis  
degraded by NaOH treatment and cDNA was ampli  
reaction. The PCR products were digested with  
cloned into DraIII- digested pME18S-FL3 vecto  
obtained cDNA vectors were used for transfor  
competent cells E. coli Top10F' by electropor  
The cDNA libraries constructed by this method  
full-length enriched cDNA library."

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

0

QY

1

AGCAATCAGTCTCTGGGAGCATATCCCTCTGCGATGGTCAGGTCAATAA

86

AGCAATCAGTCTCTGGGAGTGCACCAAGCCCTCCATCTCCAGCAACAA

QY

61

CGCTGGAGGACAAAGGAT 77

Db

146

CGCTGGAGGACAAATGAT 162

RESULT 26

AA055605

LOCUS

DEFINITION

IMAGE:510405 5' similar to gb:M29540 CARCINOEMBRYONIC

423 bp mRNA linear ES

z174h11.r1 Stratagene colon (#937204) Homo sapiens cDN

CURSOR (HUMAN); mRNA sequence.  
 15605  
 15605.1 GI:1547962  
 sapiens (human)  
 sapiens  
 alyota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 (bases 1 to 423)  
 lier.L., Lennon.G., Becker.M., Bonaldo.M.F., Chiapelli.B.,  
 soe.S., Dietrich.N., DuBuque.T., Favello.A., Gish.W.,  
 ins.M., Hultman.M., Kucaba.T., Lacy.M., Le.M., Le.N.,  
 lis.F., Moore.B., Morris.M., Parsons.J., Prange.C., Rifkin.L.,  
 lling.T., Schellenberg.K., Soares.M.B., Tan.F., Thierry-Mieg.J.,  
 askis.E., Underwood.K., Wohlmann.P., Waterston.K., Wilson.R.  
 Marra.M.  
 ation and analysis of 280,000 human expressed sequence tags  
 me Res. 6 (9), 807-828 (1996)  
 14478  
 9549  
 -act: Wilson RK  
 ington University School of Medicine  
 4 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 : 314 286 1800  
 : 314 286 1810  
 11: est@watson.wustl.edu  
 s clone is available royalty-free through LMLL; contact the  
 3E Consortium (info@image.llnl.gov) for further information.  
 primer: -28M13 rev2 from Amersham  
 n quality sequence stop: 376.  
 Location/Qualifiers  
 1..423  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:3813738"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:510405"  
 /tissue="tumor"  
 /cell\_line="T84 carcinoma cell line"  
 /lab\_host="SOLR cells (kanamycin resistant)"  
 /clone\_lib="Stratagene colon (#937204)"  
 /note="Organ: colon; Vector: pBluescript SK-; Site 1:  
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:  
 oligo dt. T-84 colonic epithelial cell line. Average  
 insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor  
 sequence: 5' GAATCGGACGAG 3' -3' adaptor sequence: 5'  
 CTCGAGTTTATTTT 3'"

54.0%; Score 41.6; DB 9; Length 423;  
 milarity 79.2%; Pred.No.0.011;  
 Conservative 0; Mismatches 15; Indels 1; Gaps 1;  
 GACAAATCACAGTCTCTGGGAGCATCATCTCTGCATGTCAGGTCATACCTCCAAAC 60  
 |||||  
 GACAAATCACAGTCTCTGGGAGTCCCAAGCCCTCCAT-CTCCAGCAACACTCCAAAC 102  
 |||||  
 CGTGGAGGACAGGAT 77  
 |||||  
 CGTGGAGGACAGGAT 119

166576  
 177560 Pancreas tumor III Homo sapiens cDNA 5' end similar to  
 :cinoembryonic antigen (GB:M17303), mRNA sequence.  
 166576  
 166576.1 GI:2018915  
 sapiens (human)  
 no sapiens  
 alyota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Ho  
 1 (bases 1 to 312)  
 Adams.M.D., Kerlavage.A.R., Fleischmann.R.D., Fuldner.R  
 Bult.C.J., Lee.N.H., Kirkness.E.F., Weinstock.K.G., Goc  
 White.O., Sutton.G., Blake.J.A., Brandon.R.C., Man-Wai,  
 Clayton.R.A., Cline.T.R., Cotton.M.D., Earle-Hughes.J.,  
 Fitzgerald.L.M., Fitzhugh.W.M., Fritchman.J.L., Geoghag  
 Glodek.A., Gnehm.C.L., Hanna.M.C., Hedblom.E., Hinkle.P  
 Kelley.J.M., Kelley.J.C., Liu.J.-L., Marnaros.S.M., Mer  
 Moreno-Palauques.R.F., McDonald.L.A., Nguyen.P.T., Pell  
 Phillips.C.A., Ryder.S.B., Scott.J.L., Saudak.D.M., Shi  
 Small.K.V., Spriggs.T.A., Uterback.T.R., Weidman.J.F.,  
 Bednarik.D.P., Cao.L., Cepeda.M.A., Coleman.T.A., Colli  
 Dimke.D., Feng.D.-F., Ferrie.A., Fischer.C., Hastings.G  
 He.W.W., Hu.J.S., Greene.J.M., Gruber.J., Hudson.P., Ki  
 Kozak.D.L., Kunsch.C., Hungjun.J., Li.H., Meisner.P.S.  
 Raymond.L., Wei.Y.F., Wang.J., Xu.C., Yu.G.L., Ruben.S.  
 Dillion.P.J., Fannon.M.R., Rosen.C.A., Haseltine.W.A.,  
 Fraser.C.M. and Venter.J.C.  
 Initial assessment of human gene diversity and expressi  
 based upon 83 million nucleotides of cDNA sequence  
 Nature 377 (6547 Suppl), 3-174 (1995)  
 96026280  
 7566098  
 Other ESTs: THC167592  
 Contact: Kerlavage, AR  
 Bioinformatics  
 The Institute for Genomic Research  
 9712 Medical Center Drive, Rockville, MD 20850 USA  
 Tel: 3018699056  
 Fax: 3018699423  
 Email: arkerlav@tigr.org  
 For clone availability, additional sequence and express  
 information related to this EST, please check the TIGR  
 Index (http://www.tigr.org/cdb/hgi/hgi.html)  
 Seq primer: M13 Reverse.  
 Location/Qualifiers  
 1..312  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="ATCC (inhost):171263"  
 /db\_xref="taxon:9606"  
 /dev\_stage="adult"  
 /clone\_lib="Pancreas tumor III"  
 /note="Organ: pancreas; Vector: pBluescript SK  
 EcoRI; Site 2: XhoI"

Query Match 53.2%; Score 41; DB 9; Length 312;  
 Best Local Similarity 79.2%; Pred.No.0.016;  
 Matches 61; Conservative 0; Mismatches 15; Indels 1;  
 QY 1 AGCAATCACAGTCTCTGGGAGCATCATCTCTGCATGTCAGGTCATAAC  
 |||||  
 Db 37 AGCAATCACAGTCTCTGGGAGTCCCAAGCCCTCCAT-CTCCAGCAACAAAC  
 |||||  
 QY 61 CCGTGGAGGACAGGAT 77  
 |||||  
 Db 96 CCGTGGAGGACAGGAT 112

RESULT 28  
 AAL132613  
 LOCUS  
 DEFINITION  
 z017f02.r1 Stratagene colon (#937204) Homo sapiens cDN  
 IMAGE:587163 5' similar to gb:M29540 CARCINOEMBRYONIC F  
 PRECURSOR (HUMAN); mRNA sequence.  
 AAL132613  
 ACCESSION  
 AAL132613.1 GI:1694102  
 VERSION  
 EST.  
 KEYWORDS  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Eut

malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
(bases 1 to 389)  
lier L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,  
scoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,  
kins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
lis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
ling, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
vaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.,  
Marra, M.  
eration and analysis of 280,000 human expressed sequence tags  
One Res. 6 (9), 807-828 (1996)  
44478  
9549  
tact: Wilson RK  
hington University School of Medicine  
4 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
: 314 286 1800  
: 314 286 1810  
l: est@watson.wustl.edu  
s clone is available royalty-free through LLNL; contact the  
3E Consortium (info@image.llnl.gov) for further information.  
ert Length: 1093 Std Error: 0.00  
Primer: -28X13 rev2 from Amersham  
h quality sequence stop: 324.  
Location/Qualifiers  
1. 389  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:4619465"  
/db\_xref="taxon:9606"  
/clones="IMAGE:587163"  
/tissue type="tumor"  
/cell line="T84 carcinoma cell line"  
/lab host="SOLR cells (kanamycin resistant)"  
/clone lib="Stratagene colon (#937204)"  
/note="Organ: colon; Vector: pBluescript SK-; Site 1:  
ECORI; Site 2: XhoI; Cloned unidirectionally. Primer:  
oligo dt. T-84 colonic epithelial cell line. Average  
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor  
sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5'  
CTCGAGTTTITTTTTTTTTT 3'."

53.2%; Score 41; DB 9; Length 389;  
milarity 79.2%; Pred. NO. 0.017;  
Conservative 0; Mismatches 15; Indels 1; Gaps 1;  
GACATCACAGTCTCTCGGAGCATCATCTCTGCGATGTCAGTTCATTAATCCAAAC 60  
GACATCACAGTCTCTCGGAGTCCCAAGCCCTCCAT-CTCCAGCACAATCCAAAC 114

CGTGGAGGACAAAGGAT 77  
CGTGGAGGACAAAGGAT 131

317244 287 bp mRNA linear EST 13-JAN-2001  
2-CI0157-111200-008-g05 CI0157 Homo sapiens cDNA, mRNA sequence.  
317244 817244.1 GI:12153747  
I. no sapiens (human)  
no sapiens  
karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
(bases 1 to 287)  
as Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
gai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, P.F.,  
ldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
unstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,  
Hare, W.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
MEDLINE  
PUBMED 10737800  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Ca  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM2&t2  
111200-008-g05&t3=2000-12-11&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 6  
High quality sequence stop: 287.  
Location/Qualifiers  
1. 287  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="CI0157"  
/note="Organ: colon; ins; Vector: puc18; Site 1  
Site 2: SmaI; A mini-library was made by clon  
derived from ORESTES PCR (U.S. Letters Patent  
No. 196,716 - Ludwig Institute for Cancer Res  
profiles into the pUC 18 vector. Reverse tran  
clease mRNA and cDNA amplification were perfor  
low stringency conditions."

Query Match 52.2%; Score 40.2; DB 10; Length 287;  
Local Similarity 70.1%; Pred. No. 0.027;  
Matches 54; Conservative 0; Mismatches 23; Indels 0;  
Qy 1 AGACAATCACAGTCTCTCGGAGCATCATCTCTGCGATGTCAGTTCAGTCATAAC  
Db 279 AGACAATCACAGTCTCTCGGAGTCCCAAGCCCTCCATCCATCGCACACACAA  
Qy 61 CGGTGGAGGACAAAGGAT 77  
Db 219 CGGTGGAGGACAAAGGAT 203

RESULT 30  
AA134992 327 bp mRNA linear ES  
LOCUS z023a06.k1 Stratagene colon (#937204) Homo sapiens cDNA  
IMAGE:587698 5', similar to gb:M29540 CARCINOEMBRYONIC I  
PRECURSOR (HUMAN);, mRNA sequence.  
ACCESSION AA134992  
VERSION AA134992.1 GI:1696111  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H  
1 (bases 1 to 327)  
REFERENCE Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiape  
Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish,  
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le,  
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C.,  
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thi  
Trevasakis, E., Underwood, K., Wohldmann, P., Waterston, R.  
and Marra, M.  
Generation and analysis of 280,000 human expressed seq  
Genome Res. 6 (9), 807-828 (1996)  
TITLE  
JOURNAL

4478

549

act: Wilson RK  
 ington University School of Medicine  
 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 314 286 1800  
 314 286 1810

l: est@watson.wustl.edu

clone is available royalty-free through LNL ; contact the  
 Consortium (info@image.lnl.gov) for further information.

E considered overall poor quality

it Length: 1162 Std Error: 0.00

primer: -28M13 rev2 from Amersham

quality sequence stop: 1.

Location/Qualifiers

1. .327

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:462000"

/db\_xref="taxon:9606"

/clone="IMAGE:587698"

/tissue\_type="tumor"

/cell\_line="T94 carcinoma cell line"

/lab\_host="SOLR cells (kanamycin resistant)"

/clone\_lib="Stratagene colon (#937204)"

/note="Organ: colon; Vector: pBluescript SK-; Site: 1:

EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:

Oligo df. T-84 colonic epithelial cell line. Average

insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor

sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5'

CTCGAGTTTTTTTTTTTTTTT 3'

52.2%; Score 40.2; DB 9; Length 327;

ilarity 76.6%; Pred. No. 0.028;

Conservative 0; Mismatches 17; Indels 1; Gaps 1;

TACATCACAGTCTCTGGGAGCATCATCTCTGCGATGTCAGTCACTCAAC 60

TACATCACAGTCTCTGGGAGTCCCAAGCAGCTCCATCTCCAGCACACTCAAC 114

TGTGGAGGACAAGGAT 77

TGTGGAGGACAAGGAT 131

10185 630 bp mRNA linear EST 04-SEP-2003

2686F02218 r1 686 (synonym: hlcc3) Homo sapiens cDNA clone

2686F02218 5', mRNA sequence.

10185

10185.1 GI:31916654

sapiens (human)

sapiens

aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

(bases 1 to 630)

cker, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,

, G., Han, M., and Wiemann, S.

(Bloecker, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., et al.)

ublished (2003)

act: MIPS

3

31staedter Landstr.1, D-85764 Neuherberg, Germany

s is the 5' sequence of the clone insert

ne from S. Wiemann, Molecular Genome Analysis, German Cancer

earch Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;

unched by GfR (National Research Centre for Biotechnology Ltd.,

unschweig/Germany) within the cDNA sequencing consortium of the

nan Genome Project.

s1 sequence available.

This clone (DKFZp686F02218) is available at the RZPD in  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

## source

Location/Qualifiers

1. .630

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="DKFZp686F02218"

/dev\_stage="adult"

/lab\_host="DHL0B"

/clone\_lib="686 (synonym: hlcc3)"

/note="Vector: pTriplex2; Site 1: SfIAA; Site 2:

cDNA-collection"

## ORIGIN

Query Match 52.2%; Score 40.2; DB 13; Length 630;

Best Local Similarity 70.1%; Pred. No. 0.038;

Matches 54; Conservative 0; Mismatches 23; Indels 0;

QY 1 AGACAATCACAGTCTCTGCGAGCATCATCTCTGCGATGTCAGTCACTCAAC

DB 86 AGACGTCACAGTCTCTGCGAGCTGCCAAGCCCTCCATCTCCAGCAACACT

QY 61 CCGTGGAGGACAAGGAT 77

DB 146 CCGTGGAGGACAAGGAT 162

RESULT 32

BF084626

LOCUS

BF084626 494 bp mRNA linear EST

DEFINITION RC1-CT0524-150900-022-f05 CT0524 Homo sapiens cDNA, mRNA

ACCESSION BF084626

VERSION BF084626.1 GI:10878469

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho

1 (bases 1 to 494)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., B

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Cost

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., S

Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V

O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de So

Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Ca

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=&t2=RC

900-022-f05&t3=2000-09-15&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 9

High quality sequence stop: 494.

Location/Qualifiers

1. .494

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

## FEATURES

## source

Location/Qualifiers

1. .494

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

```

/clone_lib="CT0524"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

          49.9%; Score 38.4; DB 10; Length 494;
similarity 70.8%; Pred. No. 0.13; Indels 0; Gaps 0;
Conservative 0; Mismatches 21;

TCACAGTCTCTGCGGAAGCATCATCTCTGCATGTCAGTCATTAATCTCCAAACCCGTG 65
TCACAGTCTCTGCGGAGTGCGCAAGCCCTCCATCTCCAGCAACAATCTCCAAACCCGTG 66
AGGACAGGAT 77
|||||
AGGACAGGAT 78

          49.9%; Score 38.4; DB 10; Length 494;
similarity 70.8%; Pred. No. 0.13; Indels 0; Gaps 0;
Conservative 0; Mismatches 21; Indels 0;

6 ATCACAGTCTCTGCGGAAGCATCATCTCTGCATGTCAGTCATTAATCTCCAA
104 ATCACAGTCTCTGCGGAGTGCGCAAGCCCTCCATCTCCAGCAACAATCTCCAA
QY 66 GAGGACAGGAT 77
Db 164 GAGGACAGGAT 175

RESULT 34
AA127122
LOCUS
DEFINITION
440 bp mRNA linear EST
Z187602.r1 Stratagene colon (#937204) Homo sapiens cDNA
IMAGE:511611 5' similar to gb:M29540 CARCINOEMBRYONIC A
PRECURSOR (HUMAN); mRNA sequence.
AA127122
VERSION
AA127122.1 GI:1686499
EST.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 440)
AUTHORS
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli,
Chissos, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Ruitman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., The,
Trevas, E., Underwood, K., Wohlmann, P., Waterston, R.,
and Marra, M.
Generation and analysis of 280,000 human expressed sequ
Genome Res. 6 (9), 807-828 (1996)
97044478
PUBMED
8899549
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LML; con
IMAGE Consortium (info@image.llnl.gov) for further info
Insert Length: 1245 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 361.
Location/Qualifiers
1. .440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GBB:3844106"
/db_xref="taxon:9606"
/clone="IMAGE:511611"
/tissue_type="tumor"
/cell_line="T84 carcinoma cell line"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Stratagene colon (#937204)"
/note="Organ: colon; Vector: pBluescript SK-;
EcoRI; Site 2: XhoI; Cloned unidirectionally.
Oligo dt. T-84 colonic epithelial cell line.
Insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' ad
sequence: 5' GAATTCGGCAGGAG 3' ~3' adaptor seq
CTCGAGTTTITTTTTTTTTTTT 3'"

FEATURES
source
1. .440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GBB:3844106"
/db_xref="taxon:9606"
/clone="IMAGE:511611"
/tissue_type="tumor"
/cell_line="T84 carcinoma cell line"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Stratagene colon (#937204)"
/note="Organ: colon; Vector: pBluescript SK-;
EcoRI; Site 2: XhoI; Cloned unidirectionally.
Oligo dt. T-84 colonic epithelial cell line.
Insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' ad
sequence: 5' GAATTCGGCAGGAG 3' ~3' adaptor seq
CTCGAGTTTITTTTTTTTTTTT 3'"

ORIGIN
Query Match 48.8%; Score 37.6; DB 9; Length 440;

```

```

/clone_lib="CT0524"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

          49.9%; Score 38.4; DB 10; Length 494;
similarity 70.8%; Pred. No. 0.13; Indels 0; Gaps 0;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

TCACAGTCTCTGCGGAAGCATCATCTCTGCATGTCAGTCATTAATCTCCAAACCCGTG 65
TCACAGTCTCTGCGGAGTGCGCAAGCCCTCCATCTCCAGCAACAATCTCCAAACCCGTG 66
AGGACAGGAT 77
|||||
AGGACAGGAT 78

          49.9%; Score 38.4; DB 10; Length 494;
similarity 70.8%; Pred. No. 0.13; Indels 0; Gaps 0;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

623 bp mRNA linear EST 19-MAY-2000
CT0347-160200-013-h04 CT0347 Homo sapiens cDNA, mRNA sequence.
32017
32017.1 GI:7957710
s sapiens (human)
s sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 623)
s Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
ai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
aman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
stein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
are, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
pson, A.J.
Equ sequencing of the human transcriptome with ORF expressed
ence tags
S. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
02663
37800
tact: Simpson A.J.G.
ratory of Cancer Genetics
wig Institute for Cancer Research
Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
zil
: +55-11-2704922
: +55-11-2707001
il: asimpson@ludwig.org.br
s sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
tp://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-RC3-CT0347-160
-013-h04&t3=2000-02-16&t4=1)
primer: puc 18 forward
h quality sequence start: 22
h quality sequence stop: 623.
Location/Qualifiers
1. .623
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0347"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue

```



ilarity 75.3%; Pred. No. 0.21;  
Conservative 0; Mismatches 17; Indels 2; Gaps 1;  
ACAAATCACAGTCTCTGGGAGGAGCATCATCTCTGCATGGTCAGGTCACTAATCAAC 60  
|||||  
ACAAATCACAGTCTCTGGGAGGAGCATCATCTCTGCATGGTCAGGTCACTAATCAAC 277  
GTGGAGGACAAAGGAT 77  
|||||  
GTGGAGGACAAAGGAT 294  
|||||

7770 541 bp mRNA linear EST 30-MAY-2000  
DT0045-140200-082-g08 DT0045 Homo sapiens cDNA, mRNA sequence.  
7770  
7770.1 GI:8113198  
sapiens (human)  
ryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
alia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
bases 1 to 541)  
Neto.E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
i, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
man, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
stein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
re, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
son, A.J.  
gun sequencing of the human transcriptome with ORF expressed  
ence tags  
: Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
12663  
:7800  
act: Simpson A.J.G.  
ratory of Cancer Genetics  
ig Institute for Cancer Research  
Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
:il  
+55-11-2704922  
+55-11-2707001  
i: asimpson@ludwig.org.br  
sequence was derived from the FAPESP/LICR Human Cancer Genome  
ject. This entry can be seen in the following URL  
p://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=QV3-DT0045-140  
-082-g08&t3=2000-02-14&t4=1)  
primer: puc 18 forward  
i quality sequence start: 20  
i quality sequence stop: 75.  
Location/Qualifiers  
1. .541  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="DT0045"  
/dev\_stage="Adult"  
/note="Organ: denis drash; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent applica  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

ilarity 48.1%; Score 37; DB 10; Length 541;  
Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
GACAAATCACAGTCTCTGGGAGGAGCATCATCTCTGCATGGTCAGGTCACTAATCAAC 60  
|||||  
GACAAATCACAGTCTCTGGGAGGAGCATCATCTCTGCATGGTCAGGTCACTAATCAAC 428  
|||||

QY 61 CCGTGGAGGACAAAGGAT 77  
|||||  
Db 429 CCGTGGAGGATGAGGAT 445  
|||||

RESULT 36  
AW361304 590 bp mRNA linear EST  
LOCUS RC3-CT0254-280999-011-a08 CT0254 Homo sapiens cDNA, mRNA  
DEFINITION  
ACCESSION AW361304  
VERSION AW361304.1 GI:6865954  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
bases 1 to 590)  
HCGP http://www.ludwig.org.br/ORESTES.  
The FAPESP/LICR Human Cancer Genome Project  
JOURNAL Unpublished (1999)  
COMMENT  
AUTHORS Contact: Simpson A.J.G.  
TITLE Laboratory of Cancer Genetics  
JOURNAL Ludwig Institute for Cancer Research  
COMMENT Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3&t2:  
280999-011-a08&t3=1999-09-28&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 566.  
Location/Qualifiers  
1. 590  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="CT0254"  
/note="Organ: colon; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent applica  
196,716 - Ludwig Institute for Cancer Research)  
into the pUC 18 vector. Reverse transcription  
mRNA and cDNA amplification were performed under  
stringency conditions."

Query Match 45.7%; Score 35.2; DB 10; Length 590;  
Best Local Similarity 73.8%; Pred. No. 1.4;  
Matches 59; Conservative 0; Mismatches 18; Indels 3;  
QY 1 ACACATCACAGTCTCTGGGAGGAGCATCATCTCTGCATGGT---CAGGTCAAT  
|||||  
Db 10 AGACATCACAGTCTCTGGGAGGAGCATCATCTCTGCATGGT---CAGGTCAAT  
|||||  
QY 58 AACCCCGTGGAGGACAAAGGAT 77  
|||||  
Db 70 AACCCCGTGGAGGACAAAGGAT 89  
|||||

RESULT 37  
BG950148/c  
LOCUS BG950148 296 bp mRNA linear EST  
DEFINITION CM2-CT0639-030101-681-e08 CT0639 Homo sapiens cDNA, mRNA  
ACCESSION BG950148  
VERSION BG950148.1 GI:14368319  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. (bases 1 to 296)  
 : Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Iman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Stein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., de Almeida, J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and de Souza, A.J.  
 :gun sequencing of the human transcriptome with ORF expressed  
 :ence tags  
 2. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 32663  
 37800  
 :act: Simpson A.J.G.  
 :ratry of Cancer Genetics  
 :wig Institute for Cancer Research  
 : Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 :il  
 : +55-11-2704922  
 : +55-11-2707001  
 :il: asimpson@ludwig.org.br  
 : s sequence was derived from the FAPESP/LICR Human Cancer Genome  
 :ect. This entry can be seen in the following URL  
 :P://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-CT0639-  
 :01-681-e08&t3=2001-01-03&t4=1)  
 : primer: puc 18 forward  
 : quality sequence start: 29  
 : quality sequence stop: 293.  
 : Location/Qualifiers  
 1. .296  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="CT0639"  
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:  
 SmaI; A mini-library was made by cloning products derived  
 from ORESTES PCR (U.S. Letters Patent application No.  
 196,716 - Ludwig Institute for Cancer Research) profiles  
 into the puc 18 vector. Reverse transcription of tissue  
 mRNA and cDNA amplification were performed under low  
 stringency conditions."  
 45.5%; Score 35; DB 12; Length 296;  
 similarity 56.7%; Pred. No. 1.2;  
 Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
 TACATCAGTCTCTGGGAGCATCATCTCTGCATGGTCAGTCAATCACTCCAAAC 60  
 TACGATCAGTCTATGGAGTCCCAAGCCCTCCATCTCCAGCAACATCCAAAC 106  
 GTGGAGGACCAAGG 75  
 GTGGAGGACCAAG 91  
 35665 914 bp mRNA linear EST 15-SRP-2000  
 304416F1 NIH\_MGC\_39 Homo sapiens cDNA clone IMAGE:3639027 5',  
 A sequence.  
 35665  
 35665.1 GI:10149657  
 : sapiens (human)  
 : sapiens  
 : aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 : Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 : (bases 1 to 914)  
 : -MGC http://mgi.nci.nih.gov/  
 : tional Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL COMMENT

Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCW340 row: p column: 04  
 High quality sequence stop: 705.  
 Location/Qualifiers

## FEATURES

## source

1. .914  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3639027"  
 /tissue\_type="adenocarcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_39"  
 /note="Organ: pancreas; Vector: pOTB7; Site 1:  
 Site 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites usi  
 following 5' adaptor: GGCACGAG(G). Library con  
 by Ling Hong in the laboratory of Gerald M. Rul  
 (University of California, Berkeley) using ZAP  
 synthesis kit (Stratagene) and Superscript II (T  
 Technologies)."

## ORIGIN

Query Match 44.9%; Score 34.6; DB 10; Length 914;  
 Best Local Similarity 74.0%; Pred. No. 2.6;  
 Matches 57; Conservative 0; Mismatches 19; Indels 1;  
 QY 1 AGACAATCACAGTCTCTGGGAGCATCATCTCTGCATGGTCAGTCAATCAAC  
 Db 521 AGACAATCACAGTCTCTGGGAGCTGACCAAGCCCTCCATCTCCAG-CAAAAC  
 QY 61 CGGTGGAGGACCAAGGAT 77  
 Db 580 CGGTGGAGGACCAAGGAT 596

## RESULT 39

## AA327327

## LOCUS

DEFINITION EST30641 Colon I Homo sapiens cDNA 5' end similar to  
 carcinoembryonic antigen (GB:M17303), mRNA sequence.

## ACCESSION

## AA327327

## VERSION

## AA327327.1

## KEYWORDS

## EST.

## SOURCE

## Homo sapiens (human)

## ORGANISM

## Homo sapiens

## Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut

## Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hor

## 1 (bases 1 to 237)

## ADAMS, W.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.

## Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Goc

## White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai,

## Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J.,

## Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Goodhag

## Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.

## Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Mer

## Moreno-Palanco, R.F., McDonald, L.A., Nguyen, D.T., Pell

## Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shi

## Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F.,

## Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Coll

## Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G

## He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Ki

## Kozak, D.L., Kunsch, C., Hurgun, J., Li, H., Weissner, P.S.,

## Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.J

## Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., I

ser, C.M. and Venter, J.C.  
 tial assessment of human gene diversity and expression patterns  
 ed upon 83 million nucleotides of cDNA sequence  
 ure 377 (6547 Suppl), 3-174 (1995)  
 26280  
 6098  
 er\_ESTs: THC167592  
 tact: Kerlavage, AR  
 Informatics  
 Institute for Genomic Research  
 2 Medical Center Drive, Rockville, MD 20850 USA  
 : 3018699056  
 : 3018699423  
 il: arkerlav@tigr.org  
 clone availability, additional sequence and expression  
 mation related to this EST, please check the TIGR Human Gene  
 ex (http://www.tigr.org/tdb/hgi/hgi.html)  
 primer: M13 Reverse.  
 Location/Qualifiers

1. .237  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="ATCC (inhost):128005"  
 /db\_xref="taxon:9606"  
 /dev\_stage="adult"  
 /clone\_lib="Colon I"  
 /note="Organ: colon; Vector: pBluescript SK-; Site\_1:  
 EcoRI; Site\_2: XhoI"

44.7%; Score 34.4; DB 9; Length 237;  
 ilarity 65.8%; Pred. No. 1.7;  
 Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
 ACATCAGTCTCTGCGAGGACATCTCTGCGATGTCAGTCAATCCAAACC 61  
 CGATCAGTCTATGCGAGGCCCAACCCCTTCATGACGACCAACTCCAAACC 221  
 TTGGAGGACAAAGGAT 77  
 TGGAGGATGAGGAT 237

9404 423 bp mRNA linear EST 19-MAY-2000  
 CT0215-160300-089-F08 CT0215 Homo sapiens cDNA, mRNA sequence.  
 9404  
 9404.1 GI:7944921

sapiens (human)  
 sapiens  
 ryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 alia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Bases 1 to 423)  
 Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 i, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 man, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H.,  
 stein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,  
 re, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 son, A.J.  
 gun sequencing of the human transcriptome with ORF expressed  
 ence tags  
 . Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 2663  
 7800  
 act: Simpson A.J.G.  
 ratory of Cancer Genetics  
 ig Institute for Cancer Research  
 Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 il  
 +55-11-2704922  
 +55-11-2707001

Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Ca  
 project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=II  
 300-089-F08&t3=2000-03-16&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 362.

# FEATURES

source  
 1. .423  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="CT0215"  
 /note="Organ: colon; Vector: puc18; Site 1: Sm  
 SmaI; A mini-library was made by cloning produ  
 from ORESTES PCR (U.S. Letters Patent applica  
 196,716 - Ludwig Institute for Cancer Research  
 into the pUC 18 vector. Reverse transcription  
 mRNA and cDNA amplification were performed und  
 stringency conditions."

## ORIGIN

Query Match 44.7%; Score 34.4; DB 10; Length 423;  
 Best Local Similarity 65.8%; Pred. No. 2.1;  
 Matches 50; Conservative 0; Mismatches 26; Indels 0;  
 QY 2 GACAATCAGTCTCTGCGAGGACATCTCTCTGATGTCAGTCAATACT  
 DB 228 GACGATCAGTCTATGCGAGGCCCAACCCCTTCATGACGACCAACT  
 QY 62 CGTGAGGACAAAGGAT 77  
 DB 288 CGTGAGGATAAGGAT 303

Search completed: April 13, 2004, 20:14:19  
 Job time : 1858.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

leic search, using sw model

April 13, 2004, 19:12:18 ; Search time 174.5 Seconds  
(without alignments)  
1655.373 Million cell updates/sec

US-10-090-326-24

77  
1 agacaatcacagtctctgcg.....aacccgtggagacaagat 77

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

2475585 seqs, 1875730760 residues

hits satisfying chosen parameters: 4951170

ength: 0

ength: 2000000000

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

is the number of results predicted by chance to have a  
ter than or equal to the score of the result being printed,  
diver by analysis of the total score distribution.

SUMMARIES

Query Match	Length	DB	ID	Description
58.7	1943	14	US-10-106-698-594	Sequence 594, App
58.4	1623	13	US-10-001-873-14	Sequence 14, App1
57.9	2459	9	US-09-925-301-291	Sequence 291, App
56.4	111	9	US-09-815-343-268	Sequence 268, App
56.4	111	12	US-10-097-105-268	Sequence 268, App
56.4	190	9	US-09-815-343-254	Sequence 254, App
56.4	190	12	US-10-097-105-254	Sequence 254, App
56.4	326	9	US-09-815-343-887	Sequence 887, App
56.4	326	9	US-09-815-343-1112	Sequence 1112, App
56.4	326	9	US-09-815-343-1390	Sequence 1390, App
56.4	326	12	US-10-097-105-887	Sequence 887, App
56.4	326	12	US-10-097-105-1112	Sequence 1112, App
56.4	326	12	US-10-097-105-1390	Sequence 1390, App
56.4	532	9	US-09-998-598-780	Sequence 780, App
56.4	571	9	US-09-815-343-725	Sequence 725, App

16	43.4	56.4	571	12	US-10-097-105-725	Sequence
17	43.4	56.4	572	9	US-09-815-343-1124	Sequence
18	43.4	56.4	572	12	US-10-097-105-1124	Sequence
19	43.4	56.4	2106	15	US-10-168-417A-3	Sequence
20	43.4	56.4	2220	9	US-09-756-551A-16	Sequence
21	43.4	56.4	2349	14	US-10-267-384-145	Sequence
22	43.4	56.4	2364	14	US-10-198-846-10292	Sequence
23	43.4	56.4	2434	14	US-10-267-384-144	Sequence
24	43.4	56.4	2728	10	US-09-814-353-20888	Sequence
25	43.4	56.4	2974	9	US-09-954-456-56	Sequence
26	43.4	56.4	2974	9	US-09-880-107-2317	Sequence
27	43.4	56.4	2974	12	US-10-240-425-1314	Sequence
28	43.4	56.4	2974	14	US-10-157-031-340	Sequence
29	43.4	56.4	2974	14	US-10-207-655-86	Sequence
30	43.4	56.4	2974	15	US-10-117-937-593	Sequence
31	43.4	56.4	2974	15	US-10-295-027-331	Sequence
32	41	53.2	588	10	US-09-871-161-447	Sequence
33	39	50.6	710	14	US-10-066-543-95	Sequence
34	32.8	42.6	456	9	US-09-998-598-1691	Sequence
35	32.8	42.6	472	9	US-09-736-457-1050	Sequence
36	32.8	42.6	472	9	US-09-902-941-1050	Sequence
37	32.8	42.6	472	9	US-09-849-626-1050	Sequence
38	32.8	42.6	472	14	US-10-017-754-1050	Sequence
39	32.8	42.6	472	14	US-10-113-872-1050	Sequence
40	32.8	42.6	476	9	US-09-920-300A-435	Sequence
41	32.8	42.6	476	9	US-09-998-598-234	Sequence
42	32.8	42.6	476	13	US-10-033-528-435	Sequence
43	32.8	42.6	476	14	US-10-039-926-435	Sequence
44	32.8	42.6	478	9	US-09-920-300A-678	Sequence
45	32.8	42.6	478	9	US-09-998-598-293	Sequence

ALIGNMENTS

RESULT 1

US-10-106-698-594  
; Sequence 594, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleot  
; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 594  
; LENGTH: 1943  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (117)..(1177)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-106-698-594

Query Match 58.7%; Score 45.2; DB 14; Length 1943;  
Best Local Similarity 72.7%; Pred. No. 4.7e-07;  
Matches 56; Conservative 1; Mismatches 20; Indels 0;  
QY 1 AGACAATCACAGTCTCTCGGAGCATCATCTCTGATGTGAGTCACTAACT  
DB 1155 AGACAATCACAGTCTCTCGGAGTGTGCCAAGCCCTCATCTCCAGCAACAAC  
QY 61 CCGTGGAGGACAAGAT 77  
|||||

CGTGGAGGACAGGAT 1231

4 Application US/10001873  
O. US20020160388A1  
MATION:

acina, Roberto  
Recipon, Herve  
Chen, Sei-Yu  
Sun, Yongming  
Liu, Chenghua  
Turner, Leah

ENTION: Compositions and Methods Relating to Lung Specific Genes and Pro

CE: DEX-0275

ICATION NUMBER: US/10/001,873

NG DATE: 2001-11-20

ATION NUMBER: 60/252,055

DATE: 2000-11-20

ATION NUMBER: 60/252,496

Q ID NOS: 55

tentin version 3.1

3

omo sapien

isc feature

856) ..(856)

WATION: a, c, g or t

4

ilarity 58.4%; Score 45; DB 13; Length 1623;

Conservative 74.0%; Pred. No. 5.3e-07;

0; Mismatches 20; Indels 0; Gaps 0;

3ACAATCACAGTCTCTGCGGAAGCATCTCTGCGTGGTCAAGTCAATCAATCCAAAC 60

|||||

3ACAATCACAGTCTCTGCGGAAGCATCTCTGCGTGGTCAAGTCAATCCAAAC 110

|||||

CGTGGAGGACAGGAT 77

|||||

CGTGGAGGACAGGAT 127

|||||

91

Application US/09925301

20020052308A1

WATION:

esen et al.

ENTION: Nucleic Acids, Proteins and Antibodies

CE: PA106

ICATION NUMBER: US/09/925,301

NG DATE: 2001-08-10

ATION NUMBER: PCT/US00/05882

DATE: 2000-03-08

ATION NUMBER: 60/124,270

DATE: 1999-03-12

Q ID NOS: 1694

tentin Ver. 2.0

9

omo sapiens

isc feature

3)

WATION: n equals a,t,g, or c

isc feature

4)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc feature

; LOCATION: (1604)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc feature

; LOCATION: (1605)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc feature

; LOCATION: (2374)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc feature

; LOCATION: (2392)

; OTHER INFORMATION: n equals a,t,g, or c

US-09-925-301-291

Query Match

Best Local Similarity 57.9%; Score 44.6; DB 9; Length 2459;

Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACAATCACAGTCTCTGCGGAAGCATCTCTGCGTGGTCAAGTCAATCCAAAC

Db 1582 AGACAATCACAGTCTCTGCGGAAGCATCTCTGCGTGGTCAAGTCAATCCAAAC

QY 61 CGTGGAGGACAGGAT 77

Db 1642 CGTGGAGGACAGGAT 1658

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

ENTION: AND DIAGNOSIS OF COLON CANCER  
 CE: 210121.504C1  
 ICAION NUMBER: US/10/097,105  
 NG DATE: 2002-03-13  
 Q ID NOS: 1562  
 stSEQ for Windows Version 4.0

omo sapiens

68

56.4%; Score 43.4; DB 12; Length 111;  
 milarity 72.7%; Pred. No. 1.1e-06;  
 Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
 GACAATCACAGTCTCTGCGAAGCATCATCCTCTGCATGGTCAAGTCAATACTCCAAAC 60  
 GACAATCACAGTCTCTGCGAGTGTGCCAAGCCTCCATCTCCAGCAACTCCAAAC 80  
 CGTGAGGACAAGGAT 77  
 |||||  
 CGTGAGGACAAGGAT 97

54 Application US/09815343

20010055596A1

MATION:

agher, Madeleine

Xu, Jiangchun

King, Gordon E.

ENTION: COMPOSITIONS AND METHODS FOR THERAPY AND

ENTION: DIAGNOSIS OF COLON CANCER

CE: 210121.504

ICATION NUMBER: US/09/815,343

NG DATE: 2001-03-22

2 ID NOS: 1556

stSEQ for Windows Version 4.0

omo sapien

34

56.4%; Score 43.4; DB 9; Length 190;  
 milarity 72.7%; Pred. No. 1.2e-06;  
 Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
 TACAATCACAGTCTCTGCGAAGCATCATCCTCTGCATGGTCAAGTCAATACTCCAAAC 60  
 TACAATCACAGTCTCTGCGAGTGTGCCAAGCCTCCATCTCCAGCAACTCCAAAC 80  
 GTGGAGGACAAGGAT 77  
 |||||  
 GTGGAGGACAAGGAT 97

4 Application US/10097105

US20040037842A1

IATION:

agher, Madeleine Joy

ing, Gordon E.

ecrist, Heather

Farlocker, Susan L.

u, Jiangchun

ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

ENTION: AND DIAGNOSIS OF COLON CANCER

E: 210121.504C1

CAION NUMBER: US/10/097,105

CURRENT FILING DATE: 2002-03-13  
 ; NUMBER OF SEQ ID NOS: 1562  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 254  
 ; LENGTH: 190  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-097-105-254

Query Match 56.4%; Score 43.4; DB 12; Length 190;  
 Best Local Similarity 72.7%; Pred. No. 1.2e-06;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACAATCACAGTCTCTGCGAAGCATCATCCTCTGCATGGTCAAGTCAATAAC  
 |||||  
 DB 21 AGACAATCACAGTCTCTGCGAGTGTGCCAAGCCTCCATCTCCAGCAACAAC  
 |||||  
 QY 61 CGTGAGGACAAGGAT 77  
 |||||  
 DB 81 CGTGAGGACAAGGAT 97

RESULT 8

US-09-815-343-887

; Sequence 887, Application US/09815343

; Patent No. US20010055596A1

; GENERAL INFORMATION:

; APPLICANT: Meagher, Madeleine

; APPLICANT: Xu, Jiangchun

; APPLICANT: King, Gordon E.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER

; FILE REFERENCE: 210121.504

; CURRENT APPLICATION NUMBER: US/09/815,343

; CURRENT FILING DATE: 2001-03-22

; NUMBER OF SEQ ID NOS: 1556

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 887

; LENGTH: 326

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-815-343-887

Query Match 56.4%; Score 43.4; DB 9; Length 326;  
 Best Local Similarity 72.7%; Pred. No. 1.4e-06;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACAATCACAGTCTCTGCGAAGCATCATCCTCTGCATGGTCAAGTCAATAAC  
 |||||  
 DB 21 AGACAATCACAGTCTCTGCGAGTGTGCCAAGCCTCCATCTCCAGCAACAAC  
 |||||  
 QY 61 CGTGAGGACAAGGAT 77  
 |||||  
 DB 81 CGTGAGGACAAGGAT 97

RESULT 9

US-09-815-343-1112

; Sequence 1112, Application US/09815343

; Patent No. US20010055596A1

; GENERAL INFORMATION:

; APPLICANT: Meagher, Madeleine

; APPLICANT: Xu, Jiangchun

; APPLICANT: King, Gordon E.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER

; FILE REFERENCE: 210121.504

; CURRENT APPLICATION NUMBER: US/09/815,343

; CURRENT FILING DATE: 2001-03-22

; NUMBER OF SEQ ID NOS: 1556

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 1112

; LENGTH: 326

```

homo sapien
112
56.4%; Score 43.4; DB 9; Length 326;
milarity 72.7%; Pred. No. 1.4e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;
AGCAATCACAGTCTCTGCGAGCATCATCTCTGCAATGTCAGTTCAGTTCATACTCCAAAC 60
AGCAATCACAGTCTCTGCGAGCTGCCAAGCCCTCCATCTCCAGCAACACTCCAAAC 80
CCGTGGAGGACAAAGGAT 77
CCGTGGAGGACAAAGGAT 97

390
Application US/09815343
20010055596A1
MATION:
eagher, Madeleine
Xu, Jiangchun
King, Gordon E.
ENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
ENTION: DIAGNOSIS OF COLON CANCER
CE: 210121.504
ICATION NUMBER: US/09/815.343
NG DATE: 2001-03-22
Q ID NOS: 1556
stSEQ for Windows Version 4.0
0

homo sapien
390
56.4%; Score 43.4; DB 9; Length 326;
milarity 72.7%; Pred. No. 1.4e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;
GACAATCACAGTCTCTGCGAGCATCATCTCTGCAATGTCAGTTCAGTTCATACTCCAAAC 60
GACAATCACAGTCTCTGCGAGCTGCCAAGCCCTCCATCTCCAGCAACACTCCAAAC 80
CCGTGGAGGACAAAGGAT 77
CCGTGGAGGACAAAGGAT 97

87
Application US/10097105
20040037842A1
MATION:
eagher, Madeleine Joy
King, Gordon E.
Secrist, Heather
Harlocker, Susan L.
Xu, Jiangchun
ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
ENTION: AND DIAGNOSIS OF COLON CANCER
CE: 210121.504C1
ICATION NUMBER: US/10/097,105
NG DATE: 2002-03-13
Q ID NOS: 1562
stSEQ for Windows Version 4.0

homo sapiens
37
56.4%; Score 43.4; DB 12; Length 326;
Best Local Similarity 72.7%; Pred. No. 1.4e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0
Qy 1 AGCAATCACAGTCTCTGCGAGCATCATCTCTGCAATGTCAGTTCAGTTCATACTAA
Db 21 AGCAATCACAGTCTCTGCGAGCTGCCAAGCCCTCCATCTCCAGCAACAA
Qy 61 CCGTGAGGACAAAGGAT 77
Db 81 CCGTGAGGACAAAGGAT 97

RESULT 12
US-10-097-105-1112
Sequence 1112, Application US/10097105
Publication No. US20040037842A1
GENERAL INFORMATION:
APPLICANT: Meagher, Madeleine Joy
APPLICANT: King, Gordon E.
APPLICANT: Secrist, Heather
APPLICANT: Harlocker, Susan L.
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.504C1
CURRENT APPLICATION NUMBER: US/10/097,105
CURRENT FILING DATE: 2002-03-13
NUMBER OF SEQ ID NOS: 1562
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 1112
LENGTH: 326
TYPE: DNA
ORGANISM: Homo sapiens
US-10-097-105-1112

Query Match 56.4%; Score 43.4; DB 12; Length 326;
Best Local Similarity 72.7%; Pred. No. 1.4e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;
Qy 1 AGCAATCACAGTCTCTGCGAGCATCATCTCTGCAATGTCAGTTCAGTTCATACTAA
Db 21 AGCAATCACAGTCTCTGCGAGCTGCCAAGCCCTCCATCTCCAGCAACAA
Qy 61 CCGTGAGGACAAAGGAT 77
Db 81 CCGTGAGGACAAAGGAT 97

RESULT 13
US-10-097-105-1390
Sequence 1390, Application US/10097105
Publication No. US20040037842A1
GENERAL INFORMATION:
APPLICANT: Meagher, Madeleine Joy
APPLICANT: King, Gordon E.
APPLICANT: Secrist, Heather
APPLICANT: Harlocker, Susan L.
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.504C1
CURRENT APPLICATION NUMBER: US/10/097,105
CURRENT FILING DATE: 2002-03-13
NUMBER OF SEQ ID NOS: 1562
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 1390
LENGTH: 326
TYPE: DNA
ORGANISM: Homo sapiens
US-10-097-105-1390
```

Query Match	56.4%;	Score 43.4;	DB 9;	Length 571;
Best Local Similarity	72.7%;	Pred. No. 1.7e-06;		
Matches 56; Conservative	0;	Mismatches 21;	Indels	0;
QY	1	AGCAATCATCAGTCTCTGGGAAGCATCATCTCTGCATGTCAGGTCAAGTCATAAC		
Dd	21	AGCAATCATCAGTCTCTGGGAGCTGCCAAGCCCTCCATCTCTCAGCAACAAC		
QY	61	CCGTGGAGGACAAGGAT	77	
Dd	81	CCGTGGAGGACAAGGAT	97	

```

RESULT 16
US-10-097-105-725
; Sequence 725, Application US/10097105
; Publication No. US20040037842A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: King, Gordon E.
; APPLICANT: Secrist, Heather
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF COLON CANCER
; FILE REFERENCE: 210121.504C1
; CURRENT APPLICATION NUMBER: US/10/097,105
; CURRENT FILING DATE: 2002-03-13
; NUMBER OF SEQ ID NOS: 1562
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

? LENGTH: 571
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: 483, 520, 523, 535, 538, 539, 549, 558, 563, 566
? OTHER INFORMATION: n = A,T,C or G
US-10-097-105-725

```

QY 61 CCGTGGAGGACAAGGAT 77  
|||||  
Db 81 CCGTGGAGGACAAGGAT 97

```

CURRENT FILING DATE: 2001-03-22
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1124
LENGTH: 572
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:

```



```
misc feature
(1)...(572)
MATION: n = A,T,C or G
124
56.4%; Score 43.4; DB 9; Length 572;
ilarity 72.7%; Pred. No. 1.7e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;
AGCAATCACAGTCTCTCGGAGCATCATCTCTGCGATGTCAGGTCAACTCCAAAC 60
|||||
AGCAATCACAGTCTCTCGGAGCTGCCAAGCCCTCCATCTCCAGCAACTCCAAAC 80
|||||
CGTGAGGACAAAGGAT 77
CGTGAGGACAAAGGAT 97
124
Application US/10097105
O. US20040037842A1
MATION: Madeleine Joy
eagher, Gordon E.
King, Gordon E.
Secrist, Heather
Harlocker, Susan L.
Xu, Jiangchun
ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
ENTION: AND DIAGNOSIS OF COLON CANCER
CE: 210121.504C1
ICATION NUMBER: US/10/097,105
NG DATE: 2002-03-13
ID NOS: 1562
stSEQ for Windows Version 4.0
4
omo sapiens
isc feature
58, 313, 355, 378, 414, 419, 432, 456, 458, 468, 478, 483,
90, 505, 522, 526, 530, 537, 539, 540, 548, 561, 562, 567,
71
MATION: n = A,T,C or G
124
56.4%; Score 43.4; DB 12; Length 572;
ilarity 72.7%; Pred. No. 1.7e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;
ACAAATCACAGTCTCTCGGAGCATCATCTCTGCGATGTCAGGTCAACTCCAAAC 60
|||||
ACAAATCACAGTCTCTCGGAGCTGCCAAGCCCTCCATCTCCAGCAACTCCAAAC 80
|||||
GTGAGGACAAAGGAT 77
GTGAGGACAAAGGAT 97
124
Application US/10168417A
O. US20040009185A1
MATION:
stage, Peter
harber, Brian
ambhara, Suryprakash
ia, Charles Dwo Yuan
TION: Enhancing the Immune Response to an Antigen by Presensitizing with
TION: Inducing Agent Prior to Immunizing with the Inducing Agent and
E: 11014-18-US
CATION NUMBER: US/10/168,417A
```

```
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/174,587
; PRIOR FILING DATE: 2000-01-05
; PRIOR APPLICATION NUMBER: PCT/CA01/00005
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2106
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: modified CEA
US-10-168-417A-3
Query Match 56.4%; Score 43.4; DB 15; Length 2106;
Best Local Similarity 72.7%; Pred. No. 2.3e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;
QY 1 AGCAATCACAGTCTCTCGGAGCATCATCTCTGCGATGTCAGGTCAACTCCAAAC
Db 1472 AGCAATCACAGTCTCTCGGAGCTGCCAAGCCCTCCATCTCCAGCAACTCCAAAC
QY 61 CGTGAGGACAAAGGAT 77
Db 1532 CGTGAGGACAAAGGAT 1548
RESULT 20
US-09-756-551A-16
; Sequence 16, Application US/09756551A
; Patent No. US20020051768A1
; GENERAL INFORMATION:
; APPLICANT: C. Morrow et al.
; TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL
; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSER: LAHIVE & COCKFIELD
; STREET: 28 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/756,551A
; FILING DATE: 08-JAN-2001
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/376,184
; FILING DATE: 17-AUG-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/987,867
; FILING DATE: 09-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/389,459
; FILING DATE: 15-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/087,009
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauro, Peter C.
; REGISTRATION NUMBER: 32,360
; REFERENCE/DOCKET NUMBER: UAI-004CPDV2CN
; TELECOMMUNICATION INFORMATION:
```

IE: (617) 227-7400  
 FOR SEQ ID NO: 16:  
 CHARACTERISTICS:  
 2220 base pairs  
 nucleic acid  
 NESS: single  
 : linear  
 YPE: CDNA  
 : CDS  
 I: 1..2203  
 16

56.4%; Score 43.4; DB 9; Length 2220;  
 milarity 72.7%; Pred. No. 2.4e-06;  
 Conservative 21; Mismatches 0; Gaps 0;  
 GACAATCACAGTCTCTGGGAGCATCATCTCTGCGATGTCAGTGCATATCCAAAC 60  
 |||||  
 GACAATCACAGTCTCTGGGAGCTGCCAAGCCCTCCATCTCCAGCAACACTCCAAAC 1630  
 |||||  
 CCGTGAGGACCAAGGAT 77  
 |||||  
 CCGTGAGGACCAAGGAT 1647

45 Application US/10267384  
 2 US20030198623A1  
 RMATION:

ANT: Paoletti, Enzo  
 Tartaglia, James  
 Cox, William I.

OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY

PONDENCE ADDRESSES: 217

ADDRESSEE: Curtis, Morris & Safford

REET: 530 Fifth Avenue

ITY: New York

ATE: NY

UNTRY: USA

IP: 10036

3R READABLE FORM:

DIUM TYPE: Floppy disk

MPUTER: IBM PC compatible

PERATING SYSTEM: PC-DOS/MS-DOS

FTWARE: PatentIn Release #1.0, Version #1.25

PLICATION DATA:

PLICATION NUMBER: US/10/267,384

LING DATE: 09-Oct-2002

ASSIFICATION: 435

Y/AGENT INFORMATION:

ME: Frommer, William S.

GISTRATION NUMBER: 25,506

ERENCE/DOCKET NUMBER: 454310-2530

MUNICATION INFORMATION:

LEPHONE: (212) 840-3333

LEFAX: (212) 840-0712

LEX: 425066CURTMS

FOR SEQ ID NO: 145:

E CHARACTERISTICS:

NGTH: 2349 base pairs

PE: nucleic acid

RANDEDNESS: single

POLOGY: linear

E TYPE: CDNA

E DESCRIPTION: SEQ ID NO: 145:

5

ilarity 56.4%; Score 43.4; DB 14; Length 2349;  
 72.7%; Pred. No. 2.4e-06;

Matches 56; Conservative 0; Mismatches 21; Indels 0  
 QY 1 AGACAATCACAGTCTCTGGGAGCATCATCTCTGCGATGTCAGTGCATATCA  
 |||||  
 Db 1658 AGACAATCACAGTCTCTGGGAGCTGCCAAGCCCTCCATCTCCAGCAACAA  
 |||||  
 QY 61 CCGTGAGGACCAAGGAT 77  
 |||||  
 Db 1718 CCGTGAGGACCAAGGAT 1734

# RESULT 22

US-10-198-846-10292

; Sequence 10292, Application US/10198846

; Publication No. US2003009974A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Xu, Yongyao

; APPLICANT: Wang, Youzhen

; APPLICANT: Steinmann, Kathleen

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION,

; FILE REFERENCE: MRL-049

; CURRENT APPLICATION NUMBER: US/10/198,846

; CURRENT FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/306,220

; PRIOR FILING DATE: 2001-07-18

; NUMBER OF SEQ ID NOS: 14084

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10292

; LENGTH: 2364

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-198-846-10292

## Query Match

Best Local Similarity 56.4%; Score 43.4; DB 14; Length 2364;

Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACAATCACAGTCTCTGGGAGCATCATCTCTGCGATGTCAGTGCATATCA  
 |||||  
 Db 1717 AGACAATCACAGTCTCTGGGAGCTGCCAAGCCCTCCATCTCCAGCAACAC  
 |||||

QY 61 CCGTGAGGACCAAGGAT 77

Db 1777 CCGTGAGGACCAAGGAT 1793

# RESULT 23

US-10-267-384-144

; Sequence 144, Application US/10267384

; Publication No. US20030198623A1

; GENERAL INFORMATION:

; APPLICANT: Paoletti, Enzo

; APPLICANT: Tartaglia, James

; APPLICANT: Cox, William I.

; TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY

; NUMBER OF SEQUENCES: 217

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis, Morris & Safford

; STREET: 530 Fifth Avenue

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/267,384

FILING DATE: 09-Oct-2002  
 CLASSIFICATION: 435  
 KEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454310-2530  
 COMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 TELEX: 425066CURTWS  
 FOR SEQ ID NO: 144:  
 ACE CHARACTERISTICS:  
 LENGTH: 2434 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 FILE TYPE: GINA  
 ACE DESCRIPTION: SEQ ID NO: 144:  
 144  
 56.4%; Score 43.4; DB 14; Length 2434;  
 Similarity 72.7%; Pred. No. 2.4e-06;  
 Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
 AGAATACAGTCTCTCGGAGCATCATCTCTGATGTGATGTCAGTCAATCACTCCAAAC 60  
 AGAATACAGTCTCTCGGAGTGTGCCAAGCCCTCCATCTCCAGCACTCCAAAC 1714  
 CGTGGAGGACAGGAT 77  
 CGTGGAGGACAGGAT 1731  
 0888  
 8, Application US/09814353  
 O. US20030165831A1  
 MATION:  
 ee, John  
 Thompson, Pamela  
 Lillie, James  
 ENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 ENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ENTION: THERAPY OF OVARIAN CANCER  
 ICE: MRI-006B  
 ICATION NUMBER: US/09/814,353  
 NG DATE: 2001-03-21  
 ACTION NUMBER: US 60/191,031  
 DATE: 2000-03-21  
 ACTION NUMBER: US 60/207,124  
 DATE: 2000-05-25  
 ACTION NUMBER: US 60/211,940  
 DATE: 2000-06-15  
 ACTION NUMBER: US 60/216,820  
 DATE: 2000-07-07  
 ACTION NUMBER: US 60/220,661  
 DATE: 2000-07-25  
 ACTION NUMBER: US 60/257,672  
 DATE: 2000-12-21  
 Q ID NOS: 22037  
 stSEQ for Windows Version 4.0  
 88  
 89  
 omo sapiens  
 0888  
 milarity 56.4%; Score 43.4; DB 10; Length 2728;  
 Similarity 72.7%; Pred. No. 2.5e-06;  
 Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
 GACATCAGTCTCTCGGAGCATCATCTCTGATGTGATGTCAGTCAATCACTCCAAAC 60

Db 1753 AGAATACAGTCTCTCGGAGTGTGCCAAGCCCTCCATCTCCAGCAACAA  
 QY 61 CCGTGGAGGACAGGAT 77  
 Db 1813 CCGTGGAGGACAGGAT 1829  
 RESULT 25  
 US-09-954-456-56  
 ; Sequence 56, Application US/09954456  
 ; Patent No. US20020115057A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Young, Paul  
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therape  
 ; FILE OF INVENTION: Sets  
 ; FILE REFERENCE: 689290-76  
 ; CURRENT APPLICATION NUMBER: US/09/954,456  
 ; CURRENT FILING DATE: 2001-09-18  
 ; PRIOR APPLICATION NUMBER: US/60/233,617  
 ; PRIOR FILING DATE: 2000-09-18  
 ; PRIOR APPLICATION NUMBER: US/60/234,052  
 ; PRIOR FILING DATE: 2000-09-20  
 ; PRIOR APPLICATION NUMBER: US/60/234,923  
 ; PRIOR FILING DATE: 2000-09-25  
 ; PRIOR APPLICATION NUMBER: US/60/235,134  
 ; PRIOR FILING DATE: 2000-09-25  
 ; PRIOR APPLICATION NUMBER: US/60/235,637  
 ; PRIOR FILING DATE: 2000-09-26  
 ; PRIOR APPLICATION NUMBER: US/60/235,638  
 ; PRIOR FILING DATE: 2000-09-26  
 ; PRIOR APPLICATION NUMBER: US/60/235,711  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: US/60/235,720  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: US/60/235,840  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: US/60/235,863  
 ; PRIOR FILING DATE: 2000-09-27  
 ; NUMBER OF SEQ ID NOS: 2276  
 ; SOFTWARE: Patent in version 3.0  
 ; SEQ ID NO 56  
 ; LENGTH: 2974  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-954-456-56  
 Query Match 56.4%; Score 43.4; DB 9; Length 2974;  
 Best Local Similarity 72.7%; Pred. No. 2.5e-06;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0;  
 QY 1 AGAATACAGTCTCTCGGAGCATCATCTCTGATGTGATGTCAGTCAATCACTCCAAAC  
 Db 1589 AGAATACAGTCTCTCGGAGTGTGCCAAGCCCTCCATCTCCAGCAACAA  
 QY 61 CCGTGGAGGACAGGAT 77  
 Db 1649 CCGTGGAGGACAGGAT 1665  
 RESULT 26  
 US-09-880-107-2317  
 ; Sequence 2317, Application US/09880107  
 ; Patent No. US20020142981A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Horne, Darci T.  
 ; APPLICANT: Vockley, Joseph G.  
 ; APPLICANT: Scherf, Uwe  
 ; APPLICANT: Gene Logic, Inc.  
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
 ; FILE REFERENCE: 44921-5028-WO  
 ; CURRENT APPLICATION NUMBER: US/09/880,107  
 ; CURRENT FILING DATE: 2001-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/211,379

DATE: 2000-06-14  
ATION NUMBER: US 60/237,054  
DATE: 2000-10-02  
ID NOS: 3950  
tentin Ver. 2.1  
4

omo sapiens

ATION: Genbank Accession No. US20020142981A1 M29540  
317

56.4%; Score 43.4; DB 9; Length 2974;  
ilarity 72.7%; Pred. No. 2.5e-06;  
Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
GACATCACAGTCTCTCGGAGCATCATCTTCGATGTCAGTCAATACTCAAAAC 60  
GACATCACAGTCTCTCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC 1648  
CGTGAGGACAAGGAT 77  
CGTGAGGACAAGGAT 1665

114  
Application US/10240425  
US20040033502A1

ATION:  
Williams, Amanda  
Joland, Joseph F.  
ord, Reginald V.  
lvarez, Chris  
etzel, Jon C.  
cheri, Uwe  
ockley, Joseph G.  
TION: Gene Expression Profiles in Esophageal Tissue  
E: 44921-5026  
CATION NUMBER: US/10/240,425  
IG DATE: 2002-09-30  
ATION NUMBER: PCT/US01/09847  
DATE: 2001-03-28  
TION NUMBER: US 60/193,446  
DATE: 2000-03-31  
ID NOS: 1588  
entn Ver. 2.1

mo sapiens

ATION: Genbank Accession No. US20040033502A1 M29540  
14

56.4%; Score 43.4; DB 12; Length 2974;  
ilarity 72.7%; Pred. No. 2.5e-06;  
Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
ACATCACAGTCTCTCGGAGCATCATCTTCGATGTCAGTCAATACTCAAAAC 60  
ACATCACAGTCTCTCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC 1648

GTGGAGGACAAGGAT 77  
GTGGAGGACAAGGAT 1665

0  
Application US/10157031  
US2003010890A1

GENERAL INFORMATION:

APPLICANT: Baranova, A. V.  
APPLICANT: Yankovsky, N. K.  
APPLICANT: Kozlov, A. P.  
APPLICANT: Lobashev, A. V.  
APPLICANT: Krukovskaya, L. L.  
TITLE OF INVENTION: In silico screening for phenotype-associated  
FILE REFERENCE: 2760-103  
CURRENT APPLICATION NUMBER: US/10/157,031  
NUMBER OF SEQ ID NOS: 415  
CURRENT FILING DATE: 2002-05-30  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 340  
LENGTH: 2974  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-157-031-340

Query Match 56.4%; Score 43.4; DB 14; Length 2974;  
Best Local Similarity 72.7%; Pred. No. 2.5e-06;  
Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACATCACAGTCTCTCGGAGCATCATCTTCGATGTCAGTCAATACTCAAAAC  
Db 1589 AGACATCACAGTCTCTCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC  
QY 61 CGTGAGGACAAGGAT 77  
Db 1649 CGTGAGGACAAGGAT 1665

RESULT 29

US-10-207-655-86  
Sequence 86, Application US/10207655  
Publication No. US20030118592A1  
GENERAL INFORMATION:  
APPLICANT: Ledbetter, Jeffrey A.  
APPLICANT: Hayden-Ledbetter, Martha S.  
TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEIN:  
FILE REFERENCE: 390069.401C1  
CURRENT APPLICATION NUMBER: US/10/207,655  
CURRENT FILING DATE: 2002-07-25  
NUMBER OF SEQ ID NOS: 426  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 86  
LENGTH: 2974  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-207-655-86

Query Match 56.4%; Score 43.4; DB 14; Length 2974;  
Best Local Similarity 72.7%; Pred. No. 2.5e-06;  
Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACATCACAGTCTCTCGGAGCATCATCTTCGATGTCAGTCAATACT  
Db 1589 AGACATCACAGTCTCTCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC  
QY 61 CGTGAGGACAAGGAT 77  
Db 1649 CGTGAGGACAAGGAT 1665

RESULT 30

US-10-117-937-593  
Sequence 593, Application US/10117937  
Publication No. US2003022039A1  
GENERAL INFORMATION:  
APPLICANT: CTL IMMUNO THERAPIES CORP.  
APPLICANT: SIMARD, John, J. L.  
APPLICANT: DIAMOND, David, C.  
APPLICANT: LIU, Liping  
APPLICANT: XIE, Zhidong

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 331
; LENGTH: 2974
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-331

Query Match          56.4%; Score 43.4; DB 15; Length 2974
Best Local Similarity 72.7%; Pred. No. 2.5e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0

QY 1 AGACAATCACAGTCTCTCGGAAGCATCATCTCTGCATGGTCAGGTCAATAC
Db 1589 AGACAATCACAGTCTCTCGGAGCTGCCAAGCCCTCCATCTCCAGCAACA
QY 61 CCGTGGAGGACAAGGAT 77
Db 1649 CCGTGGAGGACAAGGAT 1665

RESULT 32
US-09-871-161-447
; Sequence 447, Application US/09871161
; Publication No. US20030097666A1
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/871,161
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 09/328,111
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/117,393
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 60/098,639
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 447
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(588)
; OTHER INFORMATION: n = A,T,C or G
US-09-871-161-447

Query Match          53.2%; Score 41; DB 10; Length 588;
Best Local Similarity 79.2%; Pred. No. 1.4e-05;
Matches 61; Conservative 0; Mismatches 15; Indels 1;

QY 1 AGACAATCACAGTCTCTCGGAAGCATCATCTCTGCATGGTCAGGTCAATAC
Db 314 AGACAATCACAGTCTCTCGGAGCTGCCAAGCCCTCCAT-CTCCAGCAACAAC
QY 61 CCGTGGAGGACAAGGAT 77
Db 373 CCGTGGAGGACAAGGAT 389

RESULT 33
US-10-066-543-95
; Sequence 95, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather

```

Carter, Darrick  
Fanger, Gary R.  
Smith, Carole L.  
Durham, Margarita  
Stolk, John A.

ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
ENTION: AND DIAGNOSIS OF COLON CANCER

CE: 210121.563

ICATION NUMBER: US/10/066,543

NG DATE: 2002-01-31

Q ID NOS: 3417

stSEQ for Windows Version 4.0

omo sapiens

isc feature

, 2, 551, 552, 616, 642, 647, 683, 699

MATION: n = A,T,C or G

5

50.6%; Score 39; DB 14; Length 710;

milarity 76.6%; Pred. No. 8.4e-05;

Conservative 17; Mismatches 17; Indels 1; Gaps 1;

GACAATCACAGTCTCTGCGGAGCATCATCTCTGCATGGTGCAGTCAATCCAAAC 60

|||||

GACAATCACAGTCTCTGCGGAGTGCACAGCCCTCCAT-CTCCAGCAACNNCTCCAAAC 560

|||||

CGTGGAGGACAAGGAT 77

|||||

CGTGGAGGACAAGGAT 577

591/c

. Application US/09998598

20020150922A1

ATION:

olk, John A.

Ku, Jiangchun

Chenault, Ruth A.

leagher, Madelein Joy

ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

ENTION: DIAGNOSIS OF COLON CANCER

CE: 210121.561

ICATION NUMBER: US/09/998,598

NG DATE: 2001-11-16

Q ID NOS: 2606

cixa Invention Disclosure Database

omo sapiens

591

42.6%; Score 32.8; DB 9; Length 456;

milarity 64.5%; Pred. No. 0.018;

Conservative 0; Mismatches 27; Indels 0; Gaps 0;

ACAATCACAGTCTCTGCGGAGCATCATCTCTGCATGGTGCAGTCAATCCAAAC 61

|||||

ACGATCACAGTCTATGCGAGCCCAACCCCTTCATCACAGCAACTCCAAACC 110

|||||

GTGGAGGACAAGGAT 77

|||||

GTGGAGGATGAGGAT 94

150/c

. Application US/09736457

; Patent No. US20020168637A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Lodes, Michael A.

; APPLICANT: Fanger, Gary

; APPLICANT: Vedvick, Tom

; APPLICANT: Carter, Darrick

; APPLICANT: Retter, Marc

; APPLICANT: Mannion, Jane

; APPLICANT: Fan, Liqun

; APPLICANT: Wang, Aijun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY A

; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.478C15

; CURRENT APPLICATION NUMBER: US/09/736,457

; CURRENT FILING DATE: 2000-12-13

; NUMBER OF SEQ ID NOS: 1864

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1050

; LENGTH: 472

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-736-457-1050

Query Match

Best Local Similarity 42.6%; Score 32.8; DB 9; Length 472;

Matches 49; Conservative 0; Mismatches 27; Indels 0;

QY

2 GACAATCACAGTCTCTGCGGAGCATCATCTCTGCATGGTGCAGTCAATCAACT

Db

169 GACGATCACAGTCTATGCGAGCCCAACCCCTTCATCACGAGCAACAAC

QY

62 CGTGGAGGACAAGGAT 77

Db

109 CGTGGAGGATGAGGAT 94

RESULT 36

US-09-902-941-1050/c

; Sequence 1050, Application US/09902941

; Patent No. US20020172952A1

; GENERAL INFORMATION:

; APPLICANT: Henderson, Robert A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Johnson, Jeffrey C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Marnerakis, Margarita

; APPLICANT: Carter, Darrick

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: McNabb, Andria

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.478C17

; CURRENT APPLICATION NUMBER: US/09/902,941

; CURRENT FILING DATE: 2001-07-10

; NUMBER OF SEQ ID NOS: 2002

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1050

; LENGTH: 472

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-902-941-1050

Query Match

Best Local Similarity 42.6%; Score 32.8; DB 9; Length 472;

Matches 49; Conservative 0; Mismatches 27; Indels 0;

QY

2 GACAATCACAGTCTCTGCGGAGCATCATCTCTGCATGGTGCAGTCAATCAACT

ACGATCACAGTCTATGCAGAGCCACCCAAACCCCTTCATCACAGCAACACTCCAAACC 110

3TGGAGGACAGGAT 77  
|||||  
3TGGAGGATGAGGAT 94

050/c

, Application US/09849626

3. US20020197669A1

NATION:

angur, Chaitanya

Fanger, Gary

Wang, Aijun

Wang, Tongtong

Switzer, Anne

McNeill, Patricia

Clapper, Jonathan

ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

ENTION: DIAGNOSIS OF LUNG CANCER

CE: 210121.478C16

ICATION NUMBER: US/09/849,626

NG DATE: 2001-05-03

2 ID NOS: 1926

stSEQ for Windows Version 3.0

0

omo sapien

050

42.6%; Score 32.8; DB 9; Length 472;

ilarity 64.5%; Pred. No. 0.018;

Conservative 0; Mismatches 27; Indels 0; Gaps 0;

ACAATCACAGTCTCTGCGAAGCATCATCTCTGCATGTCAGGTCATAACTCCAAAC 61

|||||

ACGATCACAGTCTATGCAGAGCCACCCAAACCCCTTCATCACAGCAACACTCCAAACC 110

3TGGAGGACAGGAT 77

|||||

3TGGAGGATGAGGAT 94

050/c

, Application US/10017754

3. US20030054363A1

NATION:

anderson, Robert A.

Wang, Tongtong

Matanabe, Yoshihiro

Johnson, Jeffrey C.

Retter, Marc W.

Warnerakis, Margarita

Carter, Darrick

Fanger, Gary R.

Vedvick, Thomas S.

Bangur, Chaitanya S.

McNabb, Andria

ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

ENTION: AND DIAGNOSIS OF LUNG CANCER

CE: 210121.478C18

ICATION NUMBER: US/10/017,754

NG DATE: 2001-10-29

Q ID NOS: 2004

stSEQ for Windows Version 4.0

0

omo sapiens

050

Query Match 42.6%; Score 32.8; DB 14; Length 472;  
Best Local Similarity 64.5%; Pred. No. 0.018;  
Matches 49; Conservative 0; Mismatches 27; Indels 0;

QY 2 GACATCACAGTCTCTGCGAAGCATCATCTCTGCATGTCAGGTCATAACT

|||||

Db 169 GAGGATCACAGTCTATGCAGAGCCACCCAAACCCCTTCATCACAGCAACACT

QY 62 CGTGGAGGACAAAGGAT 77

|||||

Db 109 CGTGGAGGATGAGGAT 94

RESULT 39

US-10-113-872-1050/c

; Sequence 1050, Application US/10113872

; Publication No. US20030170255A1

; GENERAL INFORMATION:

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Sleath, Paul R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.478C19

; CURRENT APPLICATION NUMBER: US/10/113,872

; CURRENT FILING DATE: 2002-03-28

; NUMBER OF SEQ ID NOS: 2011

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1050

; LENGTH: 472

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-113-872-1050

Query Match 42.6%; Score 32.8; DB 14; Length 472;

Best Local Similarity 64.5%; Pred. No. 0.018;

Matches 49; Conservative 0; Mismatches 27; Indels 0;

QY 2 GACATCACAGTCTCTGCGAAGCATCATCTCTGCATGTCAGGTCATAACT

|||||

Db 169 GAGGATCACAGTCTATGCAGAGCCACCCAAACCCCTTCATCACAGCAACACT

QY 62 CGTGGAGGACAAAGGAT 77

|||||

Db 109 CGTGGAGGATGAGGAT 94

RESULT 40

US-09-920-300A-435/c

; Sequence 435, Application US/09920300A

; Patent No. US20030136728A1

; GENERAL INFORMATION:

; APPLICANT: King, Gordon E.

; APPLICANT: Meagher, Madeleine Joy

; APPLICANT: Xu, Jiangchun

; APPLICANT: Secrist, Heather

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER

; FILE REFERENCE: 210121.547

; CURRENT APPLICATION NUMBER: US/09/920,300A

; CURRENT FILING DATE: 2001-07-31

; NUMBER OF SEQ ID NOS: 1789

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 435

; LENGTH: 476

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-920-300A-435

4 08:44:45 2004

us-10-090-326-24.rnpb

42.6%; Score 32.8; DB 9; Length 476;  
milarity 64.5%; Pred.No. 0.018;  
Conservative 0; Mismatches 27; Indels 0; Gaps 0;  
ACAATCACAGTCTCTGCGGAAGCATCATCTCTGTCATGGTTCAGGTCTATAACTCCAAACC 61  
|||||  
ACGATCACAGTCTATGCGAGGCCACCCAAACCCTTCATCACCAGCAACTCCAAACC 110  
|||||  
GTGGAGGACAAGGAT 77  
|||||  
GTGGAGGATGAGGAT 94  
|||||

d: April 13, 2004, 21:46:10  
5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

cleic search, using sw model

April 13, 2004, 18:19:17 ; Search time 43 Seconds  
(without alignments)  
993.749 Million cell updates/sec

US-10-090-326-24

77

1 agacatacagctctctgcg.....aacccctggaggaaggat 77

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

682709 seqs, 277475446 residues

hits satisfying chosen parameters: 1365418

length: 0

length: 2000000000

: Minimum Match 0%

: Maximum Match 100%

Listing first 45 summaries

Issued Patents NA:\*

- 1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/prodata/2/ina/PCTUS\_COMB.seq.\*
- 6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

is the number of results predicted by chance to have a  
higher than or equal to the score of the result being printed,  
derived by analysis of the total score distribution.

#### SUMMARIES

Query	Match	Length	DB	ID	Description
58.4	862	3	US-08-468-856B-4	Sequence 4, Appli	
58.4	862	3	US-08-468-859A-4	Sequence 4, Appli	
58.4	2839	3	US-08-468-856B-5	Sequence 5, Appli	
58.4	2839	3	US-08-468-859A-5	Sequence 5, Appli	
56.4	2031	1	US-08-217-299-2	Sequence 2, Appli	
56.4	2097	2	US-08-602-725-35	Sequence 35, Appli	
56.4	2220	1	US-08-389-459A-16	Sequence 16, Appli	
56.4	2220	3	US-08-987-867A-16	Sequence 16, Appli	
56.4	2349	2	US-08-184-009-145	Sequence 145, App	
56.4	2349	2	US-08-458-356-145	Sequence 145, App	
56.4	2349	3	US-08-460-736-144	Sequence 144, App	
56.4	2349	4	US-09-535-370-144	Sequence 144, App	
56.4	2434	2	US-08-184-009-144	Sequence 144, App	
56.4	2434	3	US-08-458-356-144	Sequence 144, App	
56.4	2434	3	US-08-460-736-144	Sequence 144, App	
56.4	2434	4	US-09-535-370-144	Sequence 144, App	
53.2	588	3	US-09-385-982-447	Sequence 447, App	
42.6	472	4	US-09-702-705-1050	Sequence 1050, App	
42.6	472	4	US-09-736-457-1050	Sequence 1050, App	
42.6	472	4	US-09-614-124B-1050	Sequence 1050, App	
42.6	472	4	US-09-671-325-1050	Sequence 1050, App	
42.6	534	4	US-09-702-705-1325	Sequence 1325, App	
42.6	534	4	US-09-736-457-1325	Sequence 1325, App	
42.6	534	4	US-09-614-124B-1325	Sequence 1325, App	
42.6	534	4	US-09-671-325-1325	Sequence 1325, App	
42.6	536	4	US-09-702-705-1461	Sequence 1461, App	
42.6	536	4	US-09-736-457-1461	Sequence 1461, App	

C	28	32.8	42.6	536	4	US-09-614-124B-1461	Sequenc
C	29	32.8	42.6	536	4	US-09-671-325-1461	Sequenc
	30	32.8	42.6	547	3	US-09-385-982-424	Sequenc
	31	32.8	42.6	613	3	US-09-385-982-194	Sequenc
	32	32.8	42.6	652	3	US-09-385-982-337	Sequenc
	33	32.8	42.6	656	3	US-09-385-982-316	Sequenc
	34	28	36.4	3096	4	US-09-489-039A-4015	Sequenc
	35	28	36.4	3152	1	US-07-924-028A-3	Sequenc
	36	28	36.4	3365	1	US-07-789-915A-1	Sequenc
	37	28	36.4	3365	1	US-08-005-002C-1	Sequenc
	38	28	36.4	3365	1	US-08-487-203A-1	Sequenc
	39	28	36.4	4279	4	US-09-446-402A-1	Sequenc
	40	28	36.4	4810	3	US-08-852-629-11	Sequenc
	41	28	36.4	4838	3	US-08-852-629-15	Sequenc
	42	28	36.4	6700	4	US-09-654-449-1	Sequenc
	43	28	36.4	7252	4	US-09-238-356-27	Sequenc
	44	28	36.4	7366	6	5169760-3	Patent N
C	45	28	36.4	7387	4	US-09-238-356-28	Sequenc

#### ALIGNMENTS

RESULT 1  
US-08-468-856B-4  
; Sequence 4, Application US/08468856B  
; Patent No. 6013772  
; GENERAL INFORMATION:  
; APPLICANT: Barnett, Thomas; Elting, James; Kamarc, Michael;  
; APPLICANT: Kretschmer, Axel  
; TITLE OF INVENTION: CDNAS CODING FOR MEMBERS OF THE  
; TITLE OF INVENTION: CARCINOEMERYONIC ANTIGEN FAMILY  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sprung Horn Kramer & Woods  
; STREET: 660 White Plains Road  
; CITY: Tarrytown  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10591-5144  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 Mb storage  
; COMPUTER: APPLE MACINTOSH 6500  
; OPERATING SYSTEM: SYSTEM 7.5  
; SOFTWARE: WordPerfect 3.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468,856B  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/027,974  
; FILING DATE: 08-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/760,031  
; FILING DATE: 13-SEP-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/274,107  
; FILING DATE: 21-NOV-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/207,678  
; FILING DATE: 16-JUN-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/060,031  
; FILING DATE: 19-JUN-1987  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/016,683  
; FILING DATE: 19-FEB-1987  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 06/896,361  
; FILING DATE: 13-AUG-1986  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kurt G. Briscoe  
; REGISTRATION NUMBER: 33,141

REGISTRATION NUMBER: MDI 242.10-KGB  
 REGISTRATION INFORMATION:  
 REFERENCE/DOCKET NUMBER: MDI 242.9-KGB  
 TELEPHONE: (914) 332-1700  
 TELEFAX: (914) 332-1844  
 INFORMATION FOR SEQ ID NO: 4:  
 CHARACTERISTICS:  
 LENGTH: 862 nucleotides  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

Query Match 58.4%; Score 45; DB 3; Length 862;  
 Best Local Similarity 74.0%; Pred. No. 8.7e-08;  
 Matches 57; Conservative 0; Mismatches 20; Indels 0;  
 GACAAATCACAGTCTCTGCGAAGCATCATCTCTGCATGTCAGGTGCATAACTCCAAAC 60  
 GACAAATCACAGTCTCTGCGAAGCATCATCTCTGCATGTCAGGTGCATAACTCCAAAC 591  
 CGTGAGGACAAAGGAT 77  
 CGTGAGGACAAAGGAT 608

4 application US/08468859A  
 22958

REGISTRATION:  
 Applicant: Barnett, Thomas; Elting, James; Kamarck, Michael;  
 Kretschmer, Axel

INVENTION: CDNAS CODING FOR MEMBERS OF THE  
 INVENTION: CARCINOEMBRYONIC ANTIGEN FAMILY  
 SEQUENCES: 14

SEQUENCE ADDRESS:  
 E: Sprung Horn Kramer & Woods  
 660 White Plains Road  
 Tarrytown  
 New York  
 USA

591-5144  
 HEADABLE FORM:  
 YPE: Diskette, 3.50 inch, 2.0 Mb storage  
 : APPLE MACINTOSH 6500  
 G SYSTEM: SYSTEM 7.5  
 : WordPerfect 3.5  
 PPLICATION DATA:  
 ION NUMBER: US/08/468,859A  
 ATE: 06-JUN-1995

CATION: 435  
 ICATION DATA:  
 ION NUMBER: US 08/027,974  
 ATE: 08-MAR-1993  
 ICATION DATA:  
 ION NUMBER: US 07/760,031  
 ATE: 13-SEP-1991  
 ICATION DATA:  
 ION NUMBER: US 07/274,107  
 ATE: 21-NOV-1988  
 ICATION DATA:  
 ION NUMBER: US 07/207,678  
 ATE: 16-JUN-1988  
 ICATION DATA:  
 ION NUMBER: US 07/060,031  
 ATE: 19-JUN-1987  
 ICATION DATA:  
 ION NUMBER: US 07/016,683  
 ATE: 19-FEB-1987  
 ICATION DATA:  
 ION NUMBER: US 06/896,361  
 ATE: 13-AUG-1986  
 GENT INFORMATION:

NAME: Kurt G. Briscoe  
 REGISTRATION NUMBER: 33,141  
 REFERENCE/DOCKET NUMBER: MDI 242.9-KGB  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (914) 332-1700  
 TELEFAX: (914) 332-1844  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 862 nucleotides  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

Query Match 58.4%; Score 45; DB 3; Length 862;  
 Best Local Similarity 74.0%; Pred. No. 8.7e-08;  
 Matches 57; Conservative 0; Mismatches 20; Indels 0;  
 QY 1 AGACAATCACAGTCTCTGCGAAGCATCATCTCTGCATGTCAGGTGCATAAAC  
 Db 532 AGACAATCACAGTCTCTGCGAAGCATCATCTCTGCATGTCAGGTGCATAAAC  
 QY 61 CGTGAGGACAAAGGAT 77  
 Db 592 CGTGAGGACAAAGGAT 608

RESULT 3

US-08-468-856B-5  
 ; Sequence 5, Application US/08468856B  
 ; Patent No. 6013772  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Barnett, Thomas; Elting, James; Kamarck, Michael;  
 ; APPLICANT: Kretschmer, Axel  
 ; TITLE OF INVENTION: CDNAS CODING FOR MEMBERS OF THE  
 ; TITLE OF INVENTION: CARCINOEMBRYONIC ANTIGEN FAMILY  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sprung Horn Kramer & Woods  
 ; STREET: 660 White Plains Road  
 ; CITY: Tarrytown  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10591-5144  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 Mb storage  
 ; COMPUTER: APPLE MACINTOSH 6500  
 ; OPERATING SYSTEM: SYSTEM 7.5  
 ; SOFTWARE: WordPerfect 3.5  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/468,856B  
 ; FILING DATE: 06-JUN-1995  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/027,974  
 ; FILING DATE: 08-MAR-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/760,031  
 ; FILING DATE: 13-SEP-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/274,107  
 ; FILING DATE: 21-NOV-1988  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/207,678  
 ; FILING DATE: 16-JUN-1988  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/060,031  
 ; FILING DATE: 19-JUN-1987  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/016,683  
 ; FILING DATE: 19-FEB-1987  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 06/896,361

```

ATE: 13-AUG-1986
GENT INFORMATION:
  NAME: Kurt G. Briscoe
  REGISTRATION NUMBER: 33,141
  E/DOCKET NUMBER: MDI 242.10-KGB
  ICACTION INFORMATION:
    E: (914) 332-1700
    FOR SEQ ID NO: 5:
      2839 nucleotides
      nucleic acid
      NESS: single
      : linear
      : 5
    milarity 58.4%; Score 45; DB 3; Length 2839;
    Conservativity 0; Mismatches 20; Indels 0; Gaps 0;
    GACAATCACAGTCTCTGCGGAGCATCATCTCTGATGTCAGGTGTCATACTCCAAAC 60
    GACAATCACAGTCTCTGCGGAGCATCATCTCTGATGTCAGGTGTCATACTCCAAAC 1538
    CGTGGAGGACAAGGAT 77
    CGTGGAGGACAAGGAT 1555
  Application US/08468859A
  22958
  RNATION:
    Barnett, Thomas; Elting, James; Kamarck, Michael;
    Kretschmer, Axel
  VENTION: CDNAS CODING FOR MEMBERS OF THE
  VENTION: CARCINOEMBRYONIC ANTIGEN FAMILY
  SEQUENCES: 14
  ENCE ADDRESS:
    3: Sprung Horn Kramer & Woods
    660 White Plains Road
    Arriytown
    New York
    USA
  591-5144
  ADABLE FORM:
    TYPE: Diskette, 3.50 inch, 2.0 Mb storage
    : APPLE MACINTOSH 6500
    : SYSTEM: SYSTEM 7.5
    : WordPerfect 3.5
  PLICATION DATA:
    ION NUMBER: US/08/468,859A
    ATE: 06-JUN-1995
    TATION: 435
  ICACTION DATA:
    ION NUMBER: US 08/027,974
    ATE: 08-MAR-1993
  ICACTION DATA:
    ION NUMBER: US 07/760,031
    ATE: 13-SEP-1991
  ICACTION DATA:
    ION NUMBER: US 07/274,107
    ATE: 21-NOV-1988
  ICACTION DATA:
    ION NUMBER: US 07/207,678
    ATE: 16-JUN-1988
  ICACTION DATA:
    ION NUMBER: US 07/060,031
    ATE: 19-JUN-1987
  ICACTION DATA:
    ION NUMBER: US 07/016,683
    ATE: 19-FEB-1987

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/896,361
; FILING DATE: 13-AUG-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: MDI 242.9-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2839 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-468-859A-5
Query Match 58.4%; Score 45; DB 3; Length 2839;
Best Local Similarity 74.0%; Pred. No. 1.3e-07;
Matches 57; Conservativity 0; Mismatches 20; Indels 0;
QY 1 AGACAATCACAGTCTCTGCGGAGCATCATCTCTGATGTCAGGTGTCATACTCCAAAC
Db 1479 AGACAATCACAGTCTCTGCGGAGCATCATCTCTGATGTCAGGTGTCATACTCCAAAC
QY 61 CGTGGAGGACAAGGAT 77
Db 1539 CGTGGAGGACAAGGAT 1555
RESULT 5
US-08-217-299-2
; Sequence 2, Application US/08217299
; Patent No. 5672513
; GENERAL INFORMATION:
; APPLICANT: Mach, J. P.
; APPLICANT: Pelegriin, A.
; APPLICANT: Tetskikh, A.
; TITLE OF INVENTION: Carcinoembryonic Antigen Derivatives
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: NJ
; COUNTRY: USA
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/217,299
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 93810214.2
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pokras, Bruce A.
; REGISTRATION NUMBER: 32,748
; REFERENCE/DOCKET NUMBER: RAN 4093/096
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-5801
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2031 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

```

TYPE: DNA (genomic)  
 AL: NO  
 SOURCE: SOURCE:  
 I: Homo sapiens

56.4%; Score 43.4; DB 1; Length 2031;  
 milarity 72.7%; Pred. No. 4.9e-07;  
 Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGAAGCATCATCTCTGATGTCAGGTCATACCTCCAAAC 60  
 GACAATCACAGTCTCTGCGAAGCATCATCTCTGATGTCAGGTCATACCTCCAAAC 134

CGTGGAGGACAGGAT 77  
 CGTGGAGGACAGGAT 1551

5 Application US/08602725  
 65710

RMATION:  
 BODMER, WALTER F  
 DURBIN, HELGA  
 SNARY, DAVID  
 STEWART, LORNA MD  
 YOUNG, SUSAN  
 BATES, PAUL A

INVENTION: MONOCLONAL ANTIBODIES FOR USE IN  
 NVENTION: DIAGNOSIS AND TREATMENT OF COLORECTAL CANCER  
 SEQUENCES: 40  
 ENCE ADDRESS:

E: NIXON & VANDERHYE P.C.  
 1100 NORTH GLEBE ROAD, 8TH FLOOR  
 RINGSTON  
 VA USA

201  
 EADABLE FORM:  
 YPE: Floppy disk  
 : IBM PC compatible  
 3 SYSTEM: PC-DOS/MS-DOS  
 : Patent In Release #1.0, Version #1.30  
 PPLICATION DATA:  
 ION NUMBER: US/08/602,725  
 ATE: 02-FEB-1996  
 CATION: 530  
 ION DATA:  
 ION NUMBER: PCT/GB94/01816  
 ATE: 19-AUG-1994  
 ION DATA:  
 ION NUMBER: GB 9317423  
 ATE: 21-AUG-1993  
 SENT INFORMATION:  
 ADOFF, B.J.  
 TION NUMBER: 36663  
 E/DOCKET NUMBER: 1090-8  
 ION INFORMATION:  
 E: 703-816-4091  
 703-816-4100

FOR SEQ ID NO: 35:  
 CHARACTERISTICS:  
 2097 base pairs  
 nucleic acid  
 : linear  
 YPE: cDNA  
 AL: NO  
 SOURCE: Human carcinoembryonic antigen

FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..2094  
 US-08-602-725-35

Query Match 56.4%; Score 43.4; DB 2; Length 2097;  
 Best Local Similarity 72.7%; Pred. No. 4.9e-07;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGCAATCACAGTCTCTGCGAAGCATCATCTCTGATGTCAGGTCATAAC  
 Db 1463 AGCAATCACAGTCTCTGCGAAGCATCATCTCTGATGTCAGGTCATAAC  
 QY 61 CCGTGGAGGACAGGAT 77  
 Db 1523 CCGTGGAGGACAGGAT 1539

## RESULT 7

US-08-389-459A-16  
 ; Sequence 16, Application US/08389459A  
 ; Patent No. 5817512  
 ; GENERAL INFORMATION:

APPLICANT: Morrow, Casey D. and Porter, Donna, C.  
 TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT POLIOVIRUS  
 TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
 TITLE OF INVENTION: USING SAME  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD  
 STREET: 60 STATE STREET, SUITE 510  
 CITY: BOSTON  
 STATE: MASSACHUSETTS  
 COUNTRY: USA  
 ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/389,459A  
 FILING DATE: 15-FEB-1995  
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/087,009  
 FILING DATE: 01-JUL-1993  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:

NAME: Silveri, Jean M.  
 REGISTRATION NUMBER: 39,030  
 REFERENCE/DOCKET NUMBER: UAG-004CP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 227-7400  
 TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2220 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:

NAME/KEY: CDS  
 LOCATION: 1..2203  
 US-08-389-459A-16

Query Match 56.4%; Score 43.4; DB 1; Length 2220;  
 Best Local Similarity 72.7%; Pred. No. 5e-07;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGCAATCACAGTCTCTGCGAAGCATCATCTCTGATGTCAGGTCATAAC  
 |||||

GACAAATCACAGTCTCTGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAATCCAAAC 1630

CGTGGAGGACAGGAT 77  
|||||  
CGTGGAGGACAGGAT 1647

16 Application US/08987867A  
53384

RMATION:

C. Morrow et al.  
NVENTION: ENCAPSULATED RECOMBINANT VIRAL  
NVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
NVENTION: USING SAME

SEQUENCES: 23

ENCE ADDRESS:

E: LAHIVE & COCKFIELD

28 STATE STREET

OSTON

MASSACHUSETTS

USA

109

ADABLE FORM:

PE: Floppy disk

: IBM PC compatible

: SYSTEM: PC-DOS/MS-DOS

: ASCII

PLICATION DATA:

ION NUMBER: US/08/987,867A

ATE: 09-DEC-1997

ATION: 424

ION NUMBER: US 08/087,009

ATE: 01-JUL-1993

ENT INFORMATION:

ers, Louis

ION NUMBER: 35,965

/DOCKET NUMBER: UAP-004CPDV

ICATION INFORMATION:

: (617) 227-7400

OR SEQ ID NO: 16:

ARACTERISTICS:

2220 base pairs

ucleic acid

ESS: single

linear

PE: CDNA

CDS

1..2203

6

ilarity 56.4%; Score 43.4; DB 3; Length 2220;

Conservative 72.7%; Pred. No. 5e-07;

Mismatches 21; Indels 0; Gaps 0;

AGCAATCACAGTCTCTGGAGCAAGCATCATCTCTGCATGTGTCAGGTCAATACTCCAAC 60

AGCAATCACAGTCTCTGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAATCCAAAC 1630

GTGGAGGACAAAGGAT 77

|||||

GTGGAGGACAAAGGAT 1647

5

Application US/08184009

3975

MATION:

APPLICANT: Paoletti, Enzo  
APPLICANT: Tartaglia, James  
APPLICANT: Cox, William I.  
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY  
NUMBER OF SEQUENCES: 217  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Curtis, Morris & Safford

STREET: 530 Fifth Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/184,009

FILING DATE: 19-JAN-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Frommer, William S.

REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454310-2530

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

TELEX: 425066CURTWS

INFORMATION FOR SEQ ID NO: 145:

SEQUENCE CHARACTERISTICS:

LENGTH: 2349 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-184-009-145

Query Match

Best Local Similarity 56.4%; Score 43.4; DB 2; Length 2349;

Mismatches 56; Conservative 72.7%; 0; Mismatches 21; Indels 0;

QY 1 AGCAATCACAGTCTCTGGAGCAAGCATCATCTCTGCATGTGTCAGGTCAATAAC

Db 1658 AGCAATCACAGTCTCTGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

QY 61 CGTGGAGGACAAAGGAT 77

Db 1718 CGTGGAGGACAAAGGAT 1734

RESULT 10

US-08-458-356-145

; Sequence 145, Application US/08458356

; Patent No. 5942235

; GENERAL INFORMATION:

APPLICANT: Paoletti, Enzo

APPLICANT: Tartaglia, James

APPLICANT: Cox, William I.

TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY

NUMBER OF SEQUENCES: 217

CORRESPONDENCE ADDRESS:

ADDRESSEE: Curtis, Morris & Safford

STREET: 530 Fifth Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

PLICATION DATA:  
 ION NUMBER: US/08/458,356  
 ATE: 02-JUN-1995  
 CATION: 424  
 ION INFORMATION:  
 ION NUMBER: US 08/184,009  
 ATE: 19-JAN-1994  
 GENT INFORMATION:  
 rommer, William S.  
 TION NUMBER: 25,506  
 E/DOCKET NUMBER: 454310-2530  
 ION INFORMATION:  
 E: (212) 840-3333  
 (212) 840-0712  
 425066CURTMS  
 FOR SEQ ID NO: 145:  
 CHARACTERISTICS:  
 2349 base pairs  
 nucleic acid  
 NESS: single  
 YPE: linear  
 YPE: CDNA  
 45

56.4%; Score 43.4; DB 2; Length 2349;  
 milarity 72.7%; Pred. No. 5.1e-07;  
 Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
 GCAATCACAGTCTCTGGGAGCATCATCTCTGCGAGTCCATCTCCAGCAACAAC 60  
 |||||  
 GCAATCACAGTCTCTGGGAGTCTGCGAGCTCCATCTCCAGCAACAAC 1717  
 |||||  
 CGTGGAGGACAGGAT 77  
 |||||  
 CGTGGAGGACAGGAT 1734

45  
 Application US/08460736  
 65189  
 RMATION:  
 Paoletti, Enzo  
 Tartaglia, James  
 Cox, William I.  
 NVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY  
 SEQUENCES: 217  
 ENCE ADDRESS:  
 E: Curtis, Morris & Safford  
 530 Fifth Avenue  
 ew York  
 NY  
 USA  
 036  
 EADABLE FORM:  
 YPE: Floppy disk  
 : IBM PC compatible  
 G SYSTEM: PC-DOS/MS-DOS  
 : Patent In Release #1.0, Version #1.25  
 PLICATION DATA:  
 ION NUMBER: US/08/460,736  
 ATE: 02-JUN-1995  
 CATION: 514  
 ION INFORMATION:  
 ION NUMBER: US 08/184,009  
 ATE: 19-JAN-1994  
 GENT INFORMATION:  
 rommer, William S.  
 TION NUMBER: 25,506  
 E/DOCKET NUMBER: 454310-2530  
 ION INFORMATION:  
 E: (212) 840-3333  
 (212) 840-0712

TELEX: 425066CURTMS  
 ; INFORMATION FOR SEQ ID NO: 145:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2349 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: CDNA  
 ; US-08-460-736-145  
 Query Match 56.4%; Score 43.4; DB 3; Length 2349;  
 Best Local Similarity 72.7%; Pred. No. 5.1e-07;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0;  
 QY 1 AGACAATCACAGTCTCTGGGAGCATCATCTCTGCGAGTCCATCTCCAGTGTCAATCAAC  
 |||||  
 Db 1658 AGACAATCACAGTCTCTGGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC  
 |||||  
 QY 61 CGGTGGAGGACAGGAT 77  
 |||||  
 Db 1718 CGGTGGAGGACAGGAT 1734

RESULT 12  
 US-09-535-370-145  
 ; Sequence 145, Application US/09535370  
 ; Patent No. 6537594  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Paoletti, Enzo  
 ; Tartaglia, James  
 ; Cox, William I.  
 ; TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY  
 ; NUMBER OF SEQUENCES: 217  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Curtis, Morris & Safford  
 ; STREET: 530 Fifth Avenue  
 ; CITY: New York  
 ; STATE: NY  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/535,370  
 ; FILING DATE: 24-Mar-2000  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/460,736  
 ; FILING DATE: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Frommer, William S.  
 ; REGISTRATION NUMBER: 25,506  
 ; REFERENCE/DOCKET NUMBER: 454310-2530  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 840-3333  
 ; TELEFAX: (212) 840-0712  
 ; TELEX: 425066CURTMS  
 ; INFORMATION FOR SEQ ID NO: 145:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2349 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: CDNA  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 145:  
 ; US-09-535-370-145

Query Match 56.4%; Score 43.4; DB 4; Length 2349;  
 Best Local Similarity 72.7%; Pred. No. 5.1e-07;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0;

14  
 Application US/08184009  
 13975  
 INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY  
 SEQUENCES: 217  
 INVENTOR: Paoletti, Enzo  
 COX, William I.  
 ADDRESS: 530 Fifth Avenue  
 CITY: New York  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10036

136  
 APPLICATION: US/08/184,009  
 FILING DATE: 19-JAN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454310-2530  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 TELEX: 425066CURTMS  
 INFORMATION FOR SEQ ID NO: 144:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2434 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-458-356-144

137  
 APPLICATION: US/08/184,009  
 FILING DATE: 19-JAN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454310-2530  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 TELEX: 425066CURTMS  
 INFORMATION FOR SEQ ID NO: 144:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2434 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-458-356-144

138  
 APPLICATION: US/08/184,009  
 FILING DATE: 19-JAN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454310-2530  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 TELEX: 425066CURTMS  
 INFORMATION FOR SEQ ID NO: 144:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2434 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-458-356-144

139  
 APPLICATION: US/08/184,009  
 FILING DATE: 19-JAN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454310-2530  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 TELEX: 425066CURTMS  
 INFORMATION FOR SEQ ID NO: 144:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2434 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-458-356-144

140  
 APPLICATION: US/08/184,009  
 FILING DATE: 19-JAN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454310-2530  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 TELEX: 425066CURTMS  
 INFORMATION FOR SEQ ID NO: 144:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2434 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-458-356-144

141  
 APPLICATION: US/08/184,009  
 FILING DATE: 19-JAN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454310-2530  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 TELEX: 425066CURTMS  
 INFORMATION FOR SEQ ID NO: 144:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2434 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-458-356-144

142  
 APPLICATION: US/08/184,009  
 FILING DATE: 19-JAN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454310-2530  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 TELEX: 425066CURTMS  
 INFORMATION FOR SEQ ID NO: 144:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2434 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-458-356-144

143  
 APPLICATION: US/08/184,009  
 FILING DATE: 19-JAN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454310-2530  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 TELEX: 425066CURTMS  
 INFORMATION FOR SEQ ID NO: 144:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2434 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-458-356-144

144  
 APPLICATION: US/08/184,009  
 FILING DATE: 19-JAN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454310-2530  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 TELEX: 425066CURTMS  
 INFORMATION FOR SEQ ID NO: 144:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2434 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-458-356-144

145  
 APPLICATION: US/08/184,009  
 FILING DATE: 19-JAN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454310-2530  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 TELEX: 425066CURTMS  
 INFORMATION FOR SEQ ID NO: 144:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2434 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-458-356-144

; PatentIn Release #1.0, Version #1.25

PUBLICATION DATA:

ION NUMBER: US/08/460,736

ATE: 02-JUN-1995

CATION: 514

ICATION DATA:

ION NUMBER: US 08/184,009

ATE: 19-JAN-1994

ION INFORMATION:

rommer, William S.

ION NUMBER: 25,506

E/DOCKET NUMBER: 454310-2530

ICATION INFORMATION:

E: (212) 840-3333

425066CURTMS

FOR SEQ ID NO: 144:

HARACTERISTICS:

2434 base pairs

nucleic acid

NESS: single

: linear

YPE: CDNA

44

56.4%; Score 43.4; DB 3; Length 2434;

ilarity 72.7%; Pred.No. 5.2e-07; Indels 0; Gaps 0;

Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGGGAGGAGCATCTCTGATGTCAGGTCATACTCCAAAC 60

GACAATCACAGTCTCTGGGAGGAGTCGCCAAGCCCTCCATCTCCAGCAACAATCCAAAC 1714

CTGTGGAGGACAAGGAT 77

CTGTGGAGGACAAGGAT 1731

44

Application US/09535370

37594

ORMATION:

ANT: Paolletti, Enzo

Tartaglia, James

Cox, William I.

OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY

OF SEQUENCES: 217

PONDENCE ADDRESS:

DRESSEE: Curtis, Morris & Safford

TREET: 530 Fifth Avenue

ITY: New York

DATE: NY

COUNTRY: USA

IP: 10036

ER READABLE FORM:

EDUM TYPE: Floppy disk

MPUTER: IBM PC compatible

PERATING SYSTEM: PC-DOS/MS-DOS

FTWARE: PatentIn Release #1.0, Version #1.25

T APPLICATION DATA:

PPLICATION NUMBER: US/09/535,370

ILING DATE: 24-Mar-2000

LASSIFICATION: <Unknown>

APPLICATION DATA:

PPLICATION NUMBER: 08/460,736

ILING DATE: <Unknown>

BY/AGENT INFORMATION:

AME: Frommer, William S.

EGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454310-2530

MMUNICATION INFORMATION:

ELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

TELEX: 425066CURTMS

INFORMATION FOR SEQ ID NO: 144:

SEQUENCE CHARACTERISTICS:

LENGTH: 2434 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 144:

US-09-535-370-144

Query Match

Best Local Similarity 56.4%; Score 43.4; DB 4; Length 2434;

Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACAATCACAGTCTCTGGGAGGAGCATCTCTGATGTCAGGTCATACTCCAAAC

Db 1655 AGACAATCACAGTCTCTGGGAGGAGTCGCCAAGCCCTCCATCTCCAGCAACAAC

QY 61 CCGTGGAGGACAAGGAT 77

Db 1715 CCGTGGAGGACAAGGAT 1731

RESULT 17

US-09-385-982-447

; Sequence 447, Application US/09385982

; Patent No. 6262334

; GENERAL INFORMATION:

; APPLICANT: ENDEGE, WILSON O., ET AL.

; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

; TITLE OF INVENTION: PRODUCTS: II

; FILE REFERENCE: CCDNA-260XX

; CURRENT APPLICATION NUMBER: US/09/385,982

; CURRENT FILING DATE: 1999-08-30

; EARLIER APPLICATION NUMBER: 09/328,111

; EARLIER FILING DATE: 1999-06-08

; EARLIER APPLICATION NUMBER: 60/117,393

; EARLIER FILING DATE: 1999-01-27

; EARLIER APPLICATION NUMBER: 60/098,639

; EARLIER FILING DATE: 1998-08-31

; NUMBER OF SEQ ID NOS: 544

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 447

; LENGTH: 588

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(588)

; OTHER INFORMATION: n = A,T,C or G

US-09-385-982-447

Query Match

Best Local Similarity 53.2%; Score 41; DB 3; Length 588;

Matches 61; Conservative 0; Mismatches 15; Indels 1;

QY 1 AGACAATCACAGTCTCTGGGAGGAGCATCTCTGATGTCAGGTCATACTCCAAAC

Db 314 AGACAATCACAGTCTCTGGGAGGAGTCGCCAAGCCCTCCATCTCCAGCAACAAC

QY 61 CCGTGGAGGACAAGGAT 77

Db 373 CCGTGGAGGACAAGGAT 389

RESULT 18

US-09-702-705-1050/c

; Sequence 1050, Application US/09702705

; Patent No. 6504010

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong



Bangur, Chaitanya S.  
Lodes, Michael A.  
Fanger, Gary  
Fedwick, Tom  
Ratter, Darrick  
Retter, Marc  
Mannion, Jane  
Fan, Liqun  
INTN: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
INTN: DIAGNOSIS OF LUNG CANCER  
E: 210121.478C14  
ICATION NUMBER: US/09/702,705  
G DATE: 2000-10-30  
ID NOS: 1833  
stSEQ for Windows Version 3.0  
)

pmo sapien  
950

42.6%; Score 32.8; DB 4; Length 472;  
ilarity 64.5%; Pred. No. 0.0033;  
Conservative 0; Mismatches 27; Indels 0; Gaps 0;

ACAATCAGAGTCTCTGCGGAGCATCATCCTCTGATGTCAGTTCATACTCCAAACC 61  
ACGATCAGAGTCTATGAGAGCCCAACCCCTTCATCACCAGCAACAATCCAAACC 110

FTGGAGGACAAGGAT 77  
|||||  
FTGGAGGATGAGGAT 94

950/c  
Application US/09736457  
94448  
ATION:

Bang, Tongtong  
Bangur, Chaitanya S.  
Lodes, Michael A.  
Fanger, Gary  
Fedwick, Tom  
Ratter, Darrick  
Retter, Marc  
Mannion, Jane  
Fan, Liqun  
Ang, Aijun  
INTN: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
INTN: DIAGNOSIS OF LUNG CANCER  
E: 210121.478C15  
ICATION NUMBER: US/09/736,457  
G DATE: 2000-12-13  
ID NOS: 1864  
stSEQ for Windows Version 3.0  
)

pmo sapien  
950

42.6%; Score 32.8; DB 4; Length 472;  
ilarity 64.5%; Pred. No. 0.0033;  
Conservative 0; Mismatches 27; Indels 0; Gaps 0;

ACAATCAGAGTCTCTGCGGAGCATCATCCTCTGATGTCAGTTCATACTCCAAACC 61  
ACGATCAGAGTCTATGAGAGCCCAACCCCTTCATCACCAGCAACAATCCAAACC 110

FTGGAGGACAAGGAT 77  
|||||  
FTGGAGGATGAGGAT 94

## RESULT 20

US-09-614-124B-1050/c  
Sequence 1050, Application US/09614124B  
Patent No. 6630574

## GENERAL INFORMATION:

APPLICANT: Wang, Tongtong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Fedwick, Tom  
APPLICANT: Carter, Darrick  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.478C9  
CURRENT APPLICATION NUMBER: US/09/614,124B  
CURRENT FILING DATE: 2001-07-11  
NUMBER OF SEQ ID NOS: 1668  
SOFTWARE: FastSEQ for Windows Version 3.0  
SEQ ID NO 1050  
LENGTH: 472  
TYPE: DNA  
ORGANISM: Homo sapien

US-09-614-124B-1050

Query Match 42.6%; Score 32.8; DB 4; Length 472;  
Best Local Similarity 64.5%; Pred. No. 0.0033;  
Matches 49; Conservative 0; Mismatches 27; Indels 0;

QY 2 GACAATCAGTCTCTGCGGAGCATCATCCTCTGATGTCAGTTCATACT  
DB 169 GACGATCAGTCTATGAGAGCCCAACCCCTTCATCACCAGCAACT  
QY 62 CGTGGAGGACAAGGAT 77  
DB 109 CGTGGAGGATGAGGAT 94

## RESULT 21

US-09-671-325-1050/c  
Sequence 1050, Application US/09671325  
Patent No. 6667154

## GENERAL INFORMATION:

APPLICANT: Wang, Tongtong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Fedwick, Tom  
APPLICANT: Carter, Darrick  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
APPLICANT: Fan, Liqun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.478C12  
CURRENT APPLICATION NUMBER: US/09/671,325  
CURRENT FILING DATE: 2000-09-26  
NUMBER OF SEQ ID NOS: 1825  
SOFTWARE: FastSEQ for Windows Version 3.0  
SEQ ID NO 1050  
LENGTH: 472  
TYPE: DNA  
ORGANISM: Homo sapien

US-09-671-325-1050

Query Match 42.6%; Score 32.8; DB 4; Length 472;  
Best Local Similarity 64.5%; Pred. No. 0.0033;  
Matches 49; Conservative 0; Mismatches 27; Indels 0;

QY 2 GACAATCAGTCTCTGCGGAGCATCATCCTCTGATGTCAGTTCATACT

|||||  
ACGATCAGTCTATGCGAGGACCCCAACCTTCATCAGCAGCAACTCCAAACC 110  
|||||

FTGGAGGACAAGGAT 77  
|||||  
FTGGAGGATGAGGAT 94  
|||||

325  
Application US/09702705  
04010

INVENTOR:  
Wang, Tongtong  
Bangur, Chaitanya S.  
Lodes, Michael A.  
Fanger, Gary  
Vedvick, Tom  
Carter, Darrick  
Retter, Marc  
Mannion, Jane  
Fan, Liqun

INTENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
INTENTION: DIAGNOSIS OF LUNG CANCER  
DE: 210121.478C14  
PUBLICATION NUMBER: US/09/702,705  
FILING DATE: 2000-10-30  
SEQ ID NOS: 1833  
FASTSEQ for Windows Version 3.0

homo sapien  
325

42.6%; Score 32.8; DB 4; Length 534;  
Similarity 64.5%; Pred. No. 0.0034;  
Conservative 0; Mismatches 27; Indels 0; Gaps 0;

ACAAATCAGTCTCTGCGGAGGACATCCCTTCATGTCAGTCAATCAACTCCAAACC 61  
|||||  
ACGATCAGTCTATGCGAGGACCCCAACCTTCATCAGCAGCAACTCCAAACC 425  
|||||

FTGGAGGACAAGGAT 77  
|||||  
FTGGAGGATGAGGAT 441  
|||||

325  
Application US/09736457  
09448

INVENTOR:  
Wang, Tongtong  
Bangur, Chaitanya S.  
Lodes, Michael A.  
Fanger, Gary  
Vedvick, Tom  
Carter, Darrick  
Retter, Marc  
Mannion, Jane  
Fan, Liqun  
Wang, Aijun

INTENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
INTENTION: DIAGNOSIS OF LUNG CANCER  
DE: 210121.478C15  
PUBLICATION NUMBER: US/09/736,457  
FILING DATE: 2000-12-13  
SEQ ID NOS: 1864  
FASTSEQ for Windows Version 3.0

ORGANISM: Homo sapien  
US-09-736-457-1325

Query Match 42.6%; Score 32.8; DB 4; Length 534;  
Best Local Similarity 64.5%; Pred. No. 0.0034;  
Matches 49; Conservative 0; Mismatches 27; Indels 0;

QY 2 GACAATCAGTCTCTGCGGAGGACATCATCTCTTCATGTCAGTCAATCAACT  
|||||  
DB 366 GACGATCAGTCTATGCGAGGACCCCAACCTTCATCAGCAGCAACT  
|||||  
QY 62 CGTGGAGGACAAGGAT 77  
|||||  
DB 426 CGTGGAGGATGAGGAT 441  
|||||

RESULT 24

US-09-614-124B-1325  
Sequence 1325, Application US/09614124B  
Patent No. 6630574

GENERAL INFORMATION:  
APPLICANT: Wang, Tongtong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Vedvick, Tom  
APPLICANT: Carter, Darrick  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
FILE REFERENCE: 210121.478C9  
CURRENT APPLICATION NUMBER: US/09/614,124B  
CURRENT FILING DATE: 2001-07-11  
NUMBER OF SEQ ID NOS: 1668  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 1325  
LENGTH: 534  
TYPE: DNA

ORGANISM: Homo sapien  
US-09-614-124B-1325

Query Match 42.6%; Score 32.8; DB 4; Length 534;  
Best Local Similarity 64.5%; Pred. No. 0.0034;  
Matches 49; Conservative 0; Mismatches 27; Indels 0;

QY 2 GACAATCAGTCTCTGCGGAGGACATCATCTCTTCATGTCAGTCAATCAACT  
|||||  
DB 366 GACGATCAGTCTATGCGAGGACCCCAACCTTCATCAGCAGCAACT  
|||||  
QY 62 CGTGGAGGACAAGGAT 77  
|||||  
DB 426 CGTGGAGGATGAGGAT 441  
|||||

RESULT 25

US-09-671-325-1325  
Sequence 1325, Application US/09671325  
Patent No. 6667154

GENERAL INFORMATION:  
APPLICANT: Wang, Tongtong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Vedvick, Tom  
APPLICANT: Carter, Darrick  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
APPLICANT: Fan, Liqun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.478C12  
CURRENT APPLICATION NUMBER: US/09/671,325

ING DATE: 2000-09-26  
 SEQ ID NOS: 1825  
 FastSEQ for Windows Version 3.0  
 25  
 4

Homo sapien  
 1325

42.6%; Score 32.8; DB 4; Length 534;  
 imilarity 64.5%; Pred. No. 0.0034;  
 ; Conservative 0; Mismatches 27; Indels 0; Gaps 0;  
 GACATCAGTCTCTGCGGAGCAGCATCTCTGCGATGTCAGTCTAATCTCAAACC 61  
 TACGATCAGTCTATGCGAGCCACCCAAACCTTTCATCACCAGCAACTCCAAACC 425

GTGGAGGACAAGGAT 77  
 |||||  
 .GTGGAGGATGAGGAT 441

461/c  
 Application US/09702705  
 04010

NATION:  
 Wang, Tongtong  
 Bangur, Chaitanya S.  
 Lodes, Michael A.  
 Fanger, Gary  
 Vedvick, Tom  
 Carter, Darrick  
 Retter, Marc  
 Mannion, Jane  
 Fan, Liqun  
 ENION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ENION: DIAGNOSIS OF LUNG CANCER  
 CE: 210121.478C14  
 ICATION NUMBER: US/09/702,705  
 NG DATE: 2000-10-30  
 Q ID NOS: 1833  
 astSEQ for Windows Version 3.0  
 1

omo sapien  
 461

42.6%; Score 32.8; DB 4; Length 536;  
 milarity 64.5%; Pred. No. 0.0034;  
 Conservative 0; Mismatches 27; Indels 0; Gaps 0;

ACATCAGTCTCTGCGGAGCAGCATCTCTGCGATGTCAGTCTAATCTCAAACC 61  
 TCGATCAGTCTATGCGAGCCACCCAAACCTTTCATCACCAGCAACTCCAAACC 111

GTGGAGGACAAGGAT 77  
 |||||  
 GTGGAGGATGAGGAT 95

461/c  
 Application US/09736457  
 09448

NATION:  
 Wang, Tongtong  
 Bangur, Chaitanya S.  
 Lodes, Michael A.  
 Fanger, Gary  
 Vedvick, Tom  
 Carter, Darrick

; APPLICANT: Retter, Marc  
 ; APPLICANT: Mannion, Jane  
 ; APPLICANT: Fan, Liqun  
 ; APPLICANT: Wang, Aijun  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY A  
 ; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
 ; FILE REFERENCE: 210121.478C15  
 ; CURRENT APPLICATION NUMBER: US/09/736,457  
 ; CURRENT FILING DATE: 2000-12-13  
 ; NUMBER OF SEQ ID NOS: 1864  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO 1461  
 ; LENGTH: 536  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 US-09-736-457-1461

Query Match 42.6%; Score 32.8; DB 4; Length 536;  
 Best Local Similarity 64.5%; Pred. No. 0.0034;  
 Matches 49; Conservative 0; Mismatches 27; Indels 0

QY 2 GACATCAGTCTCTGCGGAGCAGCATCTCTGCGATGTCAGTCTAATCTCAAACC  
 |||||  
 Db 170 GACATCAGTCTCTGCGAGCCACCCAAACCTTTCATCACCAGCAACTCCAAACC  
 |||||  
 QY 62 GTGGAGGACAAGGAT 77  
 |||||  
 Db 110 GTGGAGGATGAGGAT 95

# RESULT 28

US-09-614-124B-1461/c  
 ; Sequence 1461, Application US/09614124B  
 ; Patent No. 6630574

GENERAL INFORMATION:  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Bangur, Chaitanya S.  
 ; APPLICANT: Lodes, Michael A.  
 ; APPLICANT: Fanger, Gary  
 ; APPLICANT: Vedvick, Tom  
 ; APPLICANT: Carter, Darrick  
 ; APPLICANT: Retter, Marc  
 ; APPLICANT: Mannion, Jane  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
 ; FILE REFERENCE: 210121.478C9  
 ; CURRENT APPLICATION NUMBER: US/09/614,124B  
 ; CURRENT FILING DATE: 2001-07-11  
 ; NUMBER OF SEQ ID NOS: 1668  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO 1461  
 ; LENGTH: 536  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 US-09-614-124B-1461

Query Match 42.6%; Score 32.8; DB 4; Length 536;  
 Best Local Similarity 64.5%; Pred. No. 0.0034;  
 Matches 49; Conservative 0; Mismatches 27; Indels 0;

QY 2 GACATCAGTCTCTGCGGAGCAGCATCTCTGCGATGTCAGTCTAATCTCAAACC  
 |||||  
 Db 170 GACATCAGTCTCTGCGAGCCACCCAAACCTTTCATCACCAGCAACTCCAAACC  
 |||||  
 QY 62 GTGGAGGACAAGGAT 77  
 |||||  
 Db 110 GTGGAGGATGAGGAT 95

# RESULT 29

US-09-671-325-1461/c  
 ; Sequence 1461, Application US/09671325  
 ; Patent No. 6667154

## RMATION:

Wang, Tongtong  
Bangur, Chaitanya S.  
Lodes, Michael A.  
Fanger, Gary  
Vedvick, Tom  
Carter, Darrick  
Retter, Marc  
Mannion, Jane  
Fan, Liqun

## VENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

## INVENTION: DIAGNOSIS OF LUNG CANCER

INCE: 210121.478C12

PLICATION NUMBER: US/09/671,325

ING DATE: 2000-09-26

SEQ ID NOS: 1825

FastSeq for Windows Version 3.0

61

6

Homo sapien

1461

42.6%; Score 32.8; DB 4; Length 536;

Similarity 64.5%; Pred. No. 0.0034;

; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAGCATCATCTCTGATGGTCAAGTCATACTCCAAACC 61

GAGATCACAGTCTATGCGAGGAGCCCAACCCCTTCATCACAGCAAACTCCAACC 111

CGTGGAGGACAAGGAT 77

CGTGGAGGATGAGGAT 95

424

, Application US/09385982

262334

RMATION:

ENDEGE, WILSON O., ET AL.

## VENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

VENTION: PRODUCTS: II

NCE: CCDNA-260XX

LICATION NUMBER: US/09/385,982

ING DATE: 1999-08-30

LICATION NUMBER: 09/328,111

ING DATE: 1999-06-08

LICATION NUMBER: 60/117,393

ING DATE: 1999-01-27

LICATION NUMBER: 60/098,639

ING DATE: 1998-08-31

SEQ ID NOS: 544

FastSeq for Windows Version 3.0

4

7

Homo sapiens

misc\_feature

(1)...(547)

RMATION: n = A,T,C or G

424

42.6%; Score 32.8; DB 3; Length 547;

Similarity 64.5%; Pred. No. 0.0034;

; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

ZACAATCACAGTCTCTGCGGAGCATCATCTCTGATGGTCAAGTCATACTCCAAACC 61

ZAGATCACAGTCTATGCGAGGAGCCCAACCCCTTCATCACAGCAAACTCCAACC 196

QY 62 CGTGGAGGACAAGGAT 77  
Db 197 CGTGGAGGATGAGGAT 212

## RESULT 31

US-09-385-982-194

; Sequence 194, Application US/09385982

; Patent No. 6262334

; GENERAL INFORMATION:

; APPLICANT: ENDEGE, WILSON O., ET AL.

; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

; TITLE OF INVENTION: PRODUCTS: II

; FILE REFERENCE: CCDNA-260XX

; CURRENT APPLICATION NUMBER: US/09/385,982

; CURRENT FILING DATE: 1999-08-30

; EARLIER APPLICATION NUMBER: 09/328,111

; EARLIER FILING DATE: 1999-06-08

; EARLIER APPLICATION NUMBER: 60/117,393

; EARLIER FILING DATE: 1999-01-27

; EARLIER APPLICATION NUMBER: 60/098,639

; EARLIER FILING DATE: 1998-08-31

; NUMBER OF SEQ ID NOS: 544

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 194

; LENGTH: 613

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(613)

; OTHER INFORMATION: n = A,T,C or G

US-09-385-982-194

## Query Match

42.6%; Score 32.8; DB 3; Length 613;

Best Local Similarity 64.5%; Pred. No. 0.0036;

Matches 49; Conservative 0; Mismatches 27; Indels 0

QY 2 GACAATCACAGTCTCTGCGGAGCATCATCTCTGATGGTCAAGTCATACTCCAAACC

Db 140 GACGATCACAGTCTATGCGAGGAGCCCAACCCCTTCATCACAGCAAACTCCAACC

QY 62 CGTGGAGGACAAGGAT 77

Db 200 CGTGGAGGATGAGGAT 215

## RESULT 32

US-09-385-982-337

; Sequence 337, Application US/09385982

; Patent No. 6262334

; GENERAL INFORMATION:

; APPLICANT: ENDEGE, WILSON O., ET AL.

; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

; TITLE OF INVENTION: PRODUCTS: II

; FILE REFERENCE: CCDNA-260XX

; CURRENT APPLICATION NUMBER: US/09/385,982

; CURRENT FILING DATE: 1999-08-30

; EARLIER APPLICATION NUMBER: 09/328,111

; EARLIER FILING DATE: 1999-06-08

; EARLIER APPLICATION NUMBER: 60/117,393

; EARLIER FILING DATE: 1999-01-27

; EARLIER APPLICATION NUMBER: 60/098,639

; EARLIER FILING DATE: 1998-08-31

; NUMBER OF SEQ ID NOS: 544

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 337

; LENGTH: 652

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(652)

INATION: n = A,T,C or G  
137

42.6%; Score 32.8; DB 3; Length 652;  
milarity 64.5%; Pred. No. 0.0037;  
Conservative 0; Mismatches 27; Indels 0; Gaps 0;

ACAAATCACAGTCTCTGCGGAGCATCATCTCTGCATGTCAGGTCTACTTCAAAACC 61  
|||||  
ACGATCACAGTCTATGAGAGCCCAACCCCTTATCCACCAACAACTCCAAACC 196  
|||||

GTGGAGGACAAGGAT 77  
|||||  
GTGGAGGATGAGGAT 212

16 Application US/09385982  
62334

MATION:  
NBEGE, WILSON O., ET AL.  
ENTION: NOVEL HUMAN GENES AND GENE EXPRESSION  
ENTION: PRODUCTS: II

CE: CCDA-260XX  
ICATION NUMBER: US/09/385,982  
NG DATE: 1999-08-30  
ICATION NUMBER: 09/328,111  
NG DATE: 1999-06-08  
ICATION NUMBER: 60/117,393  
NG DATE: 1999-01-27  
ICATION NUMBER: 60/098,639  
NG DATE: 1998-08-31  
Q ID NOS: 544  
STSEQ for Windows Version 3.0

dmo sapiens

isc feature  
l)...(656)

INATION: n = A,T,C or G  
16

42.6%; Score 32.8; DB 3; Length 656;  
milarity 64.5%; Pred. No. 0.0037;  
Conservative 0; Mismatches 27; Indels 0; Gaps 0;

ACAAATCACAGTCTCTGCGGAGCATCATCTCTGCATGTCAGGTCTACTTCAAAACC 61  
|||||  
ACGATCACAGTCTATGAGAGCCCAACCCCTTATCCACCAACAACTCCAAACC 196  
|||||

GTGGAGGACAAGGAT 77  
|||||  
GTGGAGGATGAGGAT 212

1015 Application US/09489019A  
.0836

MATION:

Jary Breton et. al  
ENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
ENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
E: 2709.2004001

ICATION NUMBER: US/09/489,039A  
IG DATE: 2000-01-27  
ATION NUMBER: US 60/117,747  
DATE: 1999-01-29  
? ID NOS: 14342

LENGTH: 3096  
TYPE: DNA  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-4015  
Query Match 36.4%; Score 28; DB 4; Length 3096;  
Best Local Similarity 100.0%; Pred. No. 0.44;  
Matches 28; Conservative 0; Mismatches 0; Indels 0  
Qy 23 AGCATCATCTCTGTCATGGTCAGGTTCAT 50  
Db 1097 AGCATCATCTCTGTCATGGTCAGGTTCAT 1124

RESULT 35

US-07-924-028A-3  
Sequence 3, Application US/07924028A  
Patent No. 5470573  
GENERAL INFORMATION:  
APPLICANT: Lubitz Werner, Szostak, Michael P.  
TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCE  
NUMBER OF INVENTION: FOR THE PRODUCTION AND USE AS IMMUNOGENS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/924,028A  
FILING DATE: 30-SEP-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP91/00308  
FILING DATE: 02-FEB-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 40 05 874  
FILING DATE: 24-FEB-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5470573man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: HUBR 1027  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3152 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-07-924-028A-3

Query Match 36.4%; Score 28; DB 1; Length 3152;  
Best Local Similarity 100.0%; Pred. No. 0.45;  
Matches 28; Conservative 0; Mismatches 0; Indels 0;

Qy 23 AGCATCATCTCTGTCATGGTCAGGTTCAT 50  
Db 1088 AGCATCATCTCTGTCATGGTCAGGTTCAT 1115

RESULT 36

US-07-789-915A-1  
Sequence 1, Application US/07789915A  
Patent No. 5212058

```

FORMATION:
: Baker, Rohan T.
: Tobias, John W.
: Varshavsky, Alexander
INVENTION: Ubiquitin-Specific Proteases
SEQUENCES: 8
SENCE ADDRESS:
: Hamilton, Brook, Smith & Reynolds, P.C.
: Two Militia Drive
: Lexington
: Massachusetts
: U.S.A.
2173
READABLE FORM:
TYPE: Floppy disk
R: IBM PC compatible
NG SYSTEM: PC-DOS/MS-DOS
3: PatentIn Release #1.0, Version #1.25
PLICATION DATA:
TION NUMBER: US/07/789,915A
DATE: 19911108
AGENT INFORMATION:
: CDS
: 1..3363
: 1
36.4%; Score 28; DB 1; Length 3365;
milarity 100.0%; Pred. No. 0.46;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
AGCATCATCTCTGCATGTCAGGTCAT 50
|||||
AGCATCATCTCTGCATGTCAGGTCAT 1394

1
application US/08005002C
94818
Baker, Rohan T.
Tobias, John W.
Varshavsky, Alexander
NVENTION: Ubiquitin-Specific Proteases
SEQUENCES: 9
SENCE ADDRESS:
: Kevin M. Farrell, P.C.
: P.O. Box 999
: York Harbor
: Maine
: U.S.A.
911
READABLE FORM:
TYPE: Floppy disk
: IBM PC compatible
IG SYSTEM: PC-DOS/MS-DOS
: PatentIn Release #1.0, Version #1.25
PLICATION DATA:
TION NUMBER: US/08/005,002C

```

```

: FILING DATE: 15-JAN-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/789,915
: FILING DATE: 08-NOV-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Farrell, Kevin M.
: REGISTRATION NUMBER: 35,505
: REFERENCE/DOCKET NUMBER: MIT-5091AAZ
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 207-363-0558
: TELEFAX: 207-363-0528
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3365 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..3363
: US-08-005-002C-1

Query Match
Best Local Similarity 36.4%; Score 28; DB 1; Length 3365;
Matches 28; Conservative 0; Mismatches 0; Indels 0
Qy 23 AGCATCATCTCTGCATGTCAGGTCAT 50
|||||
Db 1367 AGCATCATCTCTGCATGTCAGGTCAT 1394

RESULT 38
US-08-487-203A-1
: Sequence 1, Application US/08487203A
: Patent No. 5683304
: GENERAL INFORMATION:
: APPLICANT: Baker, Rohan T.
: APPLICANT: Tobias, John W.
: APPLICANT: Varshavsky, Alexander
: TITLE OF INVENTION: Ubiquitin-Specific Proteases
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kevin M. Farrell, P.C.
: STREET: P.O. Box 999
: CITY: York Harbor
: STATE: Maine
: COUNTRY: U.S.A.
: ZIP: 03911
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/487,203A
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/005,002
: FILING DATE: 15-JAN-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Farrell, Kevin M.
: REGISTRATION NUMBER: 35,505
: REFERENCE/DOCKET NUMBER: MIT-5091A3Z
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 207-363-0558
: TELEFAX: 207-363-0528
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3365 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double

```

: 08:44:45 2004

US-10-090-326-24.rn1

linear

CDS  
1..3363

36.4%; Score 28; DB 1; Length 3363;  
ilarity 100.0%; Pred. No. 0.46;  
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CATCATCCTCTGCATGTCAGGTCAT 50  
|||||  
CATCATCCTCTGCATGTCAGGTCAT 1394

publication US/09446402A

3003

ack Jr., Charles A.

ATION: COMPOSITIONS AND METHODS FOR ACTIVATING

ATION: GENES OF INTEREST

E: 5722-2(35722/191928)

CATION NUMBER: US/09/446,402A

G DATE: 1999-12-20

TION NUMBER: PCT/US98/13093

DATE: 1998-06-24

TION NUMBER: 60/050,772

DATE: 1997-06-25

ID NOS: 19

tSEQ for Windows Version 4.0

tificial Sequence

ATION: Recombinant Molecule containing multiple cloning  
ATION: site, kozak sequence, lacZ gene.

sc\_feature

)...(64)

ATION: Multiple cloning site

sc\_feature

5)...(79)

ATION: Consensus sequence for the "Kozak sequence"

im\_transcript

0)...(4279)

ATION: Beta galactosidase

36.4%; Score 28; DB 4; Length 4279;  
ilarity 100.0%; Pred. No. 0.5;  
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CATCATCCTCTGCATGTCAGGTCAT 50

|||||

CATCATCCTCTGCATGTCAGGTCAT 1864

publication US/08852629

6825

MATION:

Moyer, Richard W

Li, Yi

Hall, Richard L

VENTION: ENTOMPOXVIRUS-VERTEBRATE GENE DELIVERY

VENTION: VECTOR AND METHOD

SEQUENCES: 17

NCE ADDRESS:

: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: Florida  
COUNTRY: U.S.A.  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/852,629  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Bencen, Gerard H  
REGISTRATION NUMBER: 35,746  
REFERENCE/DOCKET NUMBER: UF-184  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 407-426-7500  
TELEFAX: 407-839-8589  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4810 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-852-629-11

Query Match 36.4%; Score 28; DB 3; Length 4810;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 28; Conservative 0; Mismatches 0; Indels 0;

Qy 23 AGCATCATCCTCTGCATGTCAGGTCAT 50

Db 2088 AGCATCATCCTCTGCATGTCAGGTCAT 2115

Search completed: April 13, 2004, 20:16:04  
Job time : 44 secs

## SUMMARIES

Query Match	DB			ID	Description
	Length	DB			
100.0	77	6	ABQ82558	Abq82558	CEA mimic
58.7	1943	4	AAN83528	Aan83528	Human col
58.4	862	1	AAN81583	Aan81583	LV7 cDNA
58.4	862	2	AAQ54351	Aaq54351	Carcinoem
58.4	862	2	AAQ54351	Aaq54351	Human CEA
58.4	1623	6	AAD39107	Aad39107	Human lun
58.4	2839	1	AAN81584	Aan81584	LV7 cDNA
58.4	2839	1	AAN92780	Aan92780	cDNA sequ
58.4	2839	2	AAQ54352	Aaq54352	Carcinoem
58.4	2839	2	AAT46062	Aat46062	Carcinoem
57.9	2459	3	AAC77897	Aac77897	Human can
56.4	111	4	AAS57592	Aas57592	CDNA #268
56.4	190	4	AAS57578	Aas57578	CDNA #254
56.4	326	4	AAS58436	Aas58436	CDNA #111
56.4	326	4	AAS58714	Aas58714	CDNA #139
56.4	326	4	AAS58211	Aas58211	CDNA #887
56.4	532	6	AAS87469	Adv87469	Human col
56.4	571	4	AAS58049	Aas58049	CDNA #725
56.4	572	4	AAS58448	Aas58448	CDNA #112
56.4	2019	6	ABK36206	Abk36206	CDNA enco
56.4	2031	2	AAQ71567	Aaq71567	Carcinoem
56.4	2059	2	AAT36495	Aat36495	Immunogen
56.4	2097	2	AAQ82807	Aaq82807	Carcinoem

## ALIGNMENTS

OS Synthetic.

PN WO200270751-A1.

12-SEP-2002

04-MAR-2002: 2002WO-IIS006504

02-MAR-2001: 2001US-0273277P

XX (IVPT-) INTV PITTSBURY

XX  
BT  
Codfray ME  
tubetich TN

XX  
DT

XX  
PT Multiplex PCR method for detecting malignancies, e.g. adenocarcinoma of the esophagus comprises conducting a PCR amplification on a DNA sample and a PCR reaction mixture.

PS Claim 47: Page 80: 141nn: English

The present invention describes a multiplex polymerase chain reaction (PCR) (M1) comprising conducting PCR on a DNA sample in a reaction mixture conducted in first and second amplification stages, each stage comprising one or more PCR cycles comprising denaturing, annealing and elongating, where the elongating step may be conducted at the same temperature as the annealing step. The second amplification stage of (M1) is conducted under different reaction conditions from that of the first amplification stage of (M1) to modulate the relative rate of production of the first amplification product of (M1). (M1) comprises a first primer set and a second primer set, each primer set comprising a first and second amplification primer. Also described: (1) an oligonucleotide comprising 15-28 bp or its derivative; (2) an



otic; (3) rapid detection of a malignancy or of metastasised  
inoma of the oesophagus; or (4) a cartridge for use in an  
PCR system. (M1) is useful for detecting malignancies, e.g.  
inoma of the oesophagus. (M1) eliminates contamination and  
the time it takes to carry out a PCR reaction. The present  
represents an internal control probe, which is used in an  
from the present invention

77 BP; 23 A; 21 C; 18 G; 15 T; 0 U; 0 Other;

Similarity 100.0%; Score 77; DB 6; Length 77;

Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AGACATCACAGTCTCTCGGAGGACATCATCTCTGCTGTCAGTCACTCAAC 60  
|||||  
AGACATCACAGTCTCTCGGAGGACATCATCTCTGCTGTCAGTCACTCAAC 60

CGTGAGGACAGGAT 77

CGTGAGGACAGGAT 77

standard; cDNA; 1943 BP.

01 (first entry)

in cancer antigen encoding cDNA SEQ ID NO:584.

on cancer; colon cancer antigen; diagnosis; detection;

- carcinoma; ss.

ns.

10-A2.

01.

10; 2000WO-US026524.

19; 93US-0157137P.

19; 93US-0163280P.

MAN GENOME SCI INC.

Barash SC, Birse CE, Rosen CA;

-235357/24.

AG74097.

sids encoding 4277 human colon cancer-associated polypeptides,  
; preventing, diagnosing and/or treating colorectal cancers.

Page 2637-2638; 9803pp; English.

10: AAH37195 and AAG73514 to AAG77788 represent human colon  
associated nucleic acid molecules (N) and proteins (P), where the  
are collectively known as colon cancer antigens. The colon  
cancers have cytostatic activity and can be used in gene therapy  
production. N and P may be used in the prevention, diagnosis  
ment of diseases associated with inappropriate P expression. For  
N and P may be used to treat disorders associated with decreased  
by rectifying mutations or deletions in a patient's genome  
at the activity of P by expressing inactive proteins or to  
the patients' own production of P. Additionally, N may be used  
the colon cancer-associated Ps, by inserting the nucleic acids  
at cell and culturing the cell to express the proteins. N and P  
ad in the prevention, diagnosis and treatment of colorectal  
cancers. AAH37196 to AAH37204 and AAB77789 represent

CC sequences used in the exemplification of the present invention;  
CC Pages 666 to 682 and page 7053 of the sequence listing were m  
CC time of publication, meaning no sequences are present for SEQ  
CC to 1052, 7921 and 7922

SQ Sequence 1943 BP; 524 A; 597 C; 397 G; 422 T; 0 U; 3 Other;

Query Match 58.7%; Score 45.2; DB 4; Length 1943;  
Best Local Similarity 72.7%; Pred. No. 3.5e-06;  
Matches 56; Conservative 1; Mismatches 20; Indels 0

QY 1 AGACAATCACAGTCTCTCGGAGGACATCATCTCTGCTGTCAGTCACTCAAC  
|||||  
Db 1155 AGACAATCACAGTCTCTCGGAGGACATCATCTCTGCTGTCAGTCACTCAAC

QY 61 CCGTGGAGGACAGGAT 77

Db 1215 CCGTGGAGGACAGGAT 1231

RESULT 3

AAH81583

ID AAH81583 standard; cDNA; 862 BP.

AC AAH81583;

XX

DT 25-MAR-2003 (revised)

DT 07-DEC-1990 (first entry)

XX

DE LV7 cDNA encoding carcinoembryonic antigen-A.

XX

KW Carcinoembryonic antigen-A; LV7; colorectal cancer; antibodies;  
KW immunoassay; probes; ss.

XX Homo sapiens.

XX

EH Key Location/Qualifiers

FT CDS 3...860

FT /\*tag= a

FT /product= "carcinoembryonic antigen-A"

XX EP263933-A.

XX

PD 20-APR-1988.

XX

FF 03-AUG-1987; 87EP-00111168.

XX

PR 13-AUG-1986; 86US-00896361.

PR 19-FEB-1987; 87US-00016683.

PR 19-JUN-1987; 87US-00060031.

XX

FA (MOLE-) MOLECULAR DIAGNOSTICS INC.

XX

PI Barnett TR, Elting JJ, Kamarck ME;

XX

DR WPI; 1988-106756/16.

DR P-PSDB; AAP81224.

XX

PT New nucleic acid sequences coding for carcinoembryonic antigen  
PT expressed polypeptide(s), useful for making diagnostic and the  
PT antibodies.

PS Claim 1 (a); Page 46; 68pp; English.

XX

CC The sequence can be expressed in lambda gt11. It has an inter  
CC sequence of approx. 300 bases long and codes for a theoretical  
CC 285 amino acids long. This cDNA encodes for sequences found in  
CC human genes and it is believed that this entire family of gene  
CC for the CEA family of proteins. CEA is a marker for management  
CC colorectal cancer. The antibodies are useful for immunoassay  
CC detection of CEA, and also when bonded to radionuclides or to  
CC CEA-expressing cells. The nucleic acid can be used as hybrid  
CC probes. See also AAH81583-86. (Updated on 25-MAR-2003 to corre

us-10-090-326-24.rng

2 BP; 232 A; 282 C; 180 G; 168 T; 0 U; 0 Other;

andard; cDNA; 862 BP.

ytic antigen CEA-(a) coding sequence LV7.

•  
හ.

/note= "partial open reading frame, does not include initiation or termination codons"

; 86US-00896361.

87US-00060031.  
88US-00231741.

Kamarck ME, Elting JJ;

54712.

g 1; 39pp; English.

2 BP; 234 A; 280 C; 180 G; 168 T; 0 U; 0 Other;

58.4%; Score 45; DB 2; Length 862;  
74.0%; Pred. No. 3.3e-06;  
ilarity

Qy 1 AGACAATCAAGTCTCTGCGGAAGCATCATCTCTGTCATCGTTCAGGTCTATAACT  
|||||  
Db 532 AGACAATCAAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAACCT  
|||||

CGTGGAGGACCAAGGAT 77  
|||||  
CGTGGAGGACCAAGGAT 608

standard; cDNA; 1623 BP.

2 (first entry)

-specific nucleic acid (LSNA) #14.

β-specific nucleic acid; LSNA; lung-specific protein; LSP;  
c; non-cancerous disease; gene therapy; transgenic animal;  
; engineered lung tissue; cytostatic; gene; ss.

ss.

Location/Qualifiers

3. .684  
/\*tag= a  
/product= "Human LSP #14"  
/transl\_except= (pos:33. .34, aa:His)  
/transl\_except= (pos:335. .336, aa:Leu)  
/note= "These codons have an apparent deletion of 1  
nucleotide base each, which alters the reading frame; CDS  
does not include start codon"  
/partial

3-A2.

2.

1; 2001WO-US045180.

0; 2000US-0252055P.

0; 2000US-0252496P.

ADEXUS INC.

Recipon H, Chen S, Sun Y, Liu C, Turner LR;

471623/50.

E24334.

pecific nucleic acids, useful for identifying, diagnosing,  
, staging, imaging, and treating lung cancer and non-cancerous  
ates in lung tissues.

age 156-157; 187pp; English.

ion relates to new lung-specific nucleic acids (LSNA) and lung-  
roteins (LSP). The lung-specific nucleic acids, polypeptides  
itions comprising them are useful: for identifying, diagnosing,  
, staging, imaging, and treating lung cancer and non-cancerous  
ates in lung tissue; for identifying lung tissue; for  
, identifying and/or designing agonists and antagonists of the  
es; in gene therapy; in producing transgenic animals and cells;  
ing engineered lung tissue for treatment and research; and as  
n an array or computer program for pattern recognition of lung  
The nucleic acids may be used as hybridisation probes to  
aracterise and quantify hybridising nucleic acids in and  
bridging nucleic acids from, both genomic and transcript-  
cleic acid samples. The present sequence is human lung-specific

623 BP; 494 A; 419 C; 306 G; 403 T; 0 U; 1 Other;

58.4%; Score 45; DB 6; Length 1623;

milarity 74.0%; Pred. No. 3.9e-06;

Matches 57; Conservative 0; Mismatches 20; Indels 0;  
QY 1 AGACAATCACAGTCTCTGGGAAGCATCATCTCTGATGGTCAGGTCAATAC  
|||||  
Db 51 AGACAATCACAGTCTCTGGGACGTCGCCAGCCCTCCATCTCCAGCAACAAC  
|||||  
QY 61 CCGTGAGGACCAAGGAT 77  
|||||  
Db 111 CCGTGAGGACCAAGGAT 127  
|||||

RESULT 7

AAN81584

ID AAN81584 standard; DNA; 2839 BP.

XX

AC AAN81584;

XX

DT 25-MAR-2003 (revised)

DT

07-DEC-1990 (first entry)

XX

DE 1LV7 cDNA encoding carcinoembryonic antigen-B.

XX

KW Carcinoembryonic antigen-B; 1LV7; colorectal cancer; antibody;

KW

immunoassay; probes; ss.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT

CDS 5. .2110

FT

/\*tag= a

FT

/product= "carcinoembryonic antigen-B"

XX

PN EP263933-A.

XX

PD 20-APR-1988.

XX

PF 03-AUG-1987; 87EP-00111168.

XX

PR 13-AUG-1986; 86US-00896361.

PR

19-FEB-1987; 87US-00016683.

PR

19-JUN-1987; 87US-00060031.

XX

PA (MOLE-) MOLECULAR DIAGNOSTICS INC.

XX

PI Barnett TR, Elting JJ, Kamack ME;

XX

DR WPI; 1988-106756/16.

DR

P-PSDB; AAP81222.

XX

PT New nucleic acid sequences coding for carcinoembryonic antigen

PT

expressed polypeptide(s), useful for making diagnostic and the

PT

antibodies.

XX

PS Claim 1 (b); Page 46; 68pp; English.

XX

CC CEA is a marker for management of colorectal cancer. The antib

CC

useful for immunoassay and in vivo detection of CEA, and also

CC

to radionuclides or toxins to kill CEA-expressing cells. The

CC

acid can be used as hybridisation probes. See also AAN81583-86

CC

on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 t

CC

PI field.)

XX

SQ Sequence 2839 BP; 793 A; 815 C; 567 G; 664 T; 0 U; 0 Other;

Query Match

58.4%; Score 45; DB 1; Length 2839;

Best Local Similarity 74.0%; Pred. No. 4.5e-06;

Matches 57; Conservative 0; Mismatches 20; Indels 0;

QY 1 AGACAATCACAGTCTCTGGGAAGCATCATCTCTGATGGTCAGGTCAATAC

|||||

Db 1479 AGACAATCACAGTCTCTGGGACGTCGCCAGCCCTCCATCTCCAGCAACAAC

|||||

QY 61 CCGTGAGGACCAAGGAT 77

|||||  
CGTGGAGGACAGGAT 1555

standard; DNA; 2839 BP.

3 (revised)  
2 (revised)  
0 (first entry)

nce encoding CEA complete antigen 15LV7 (CEA-(b)).

ryonic antigens (CEA's); antibody (anti-CEA) production;

Location/Qualifiers  
5. .2110

9.

3; 89BP-00110096.

3; 88US-00207678.

3; 88US-00274107.

ECULAR DIAGNOSTICS INC.

Elting JJ, Kamarc ME, Kretschmer A;

372000/51.

293999.

I for members of carcinoembryonic antigen family - used to  
bes to detect antigen or in prodn. of polypeptide(s) and

Page 20-26; 65pp; English.

ice encodes the peptide sequence of the CEA-(b) antigen (pc  
is used as a nucleic acid probe to detect the presence of CEA  
a related nucleic acid in a test sample. Since different CEA's  
in different cancers the presence of a particular CEA can act  
stic tool. (Updated on 31-OCT-2002 to add missing OS field.)  
25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to  
field.)

39 BP; 790 A; 810 C; 575 G; 664 T; 0 U; 0 Other;

ilarity 58.4%; Score 45; DB 1; Length 2839;

Conservative 0; Mismatches 20; Indels 0; Gaps 0;

ACAAATCACAGTCTCTGGGAGGATCATCTCTGCATGGTCCAGTCACTCAAAAC 60  
ACAAATCACAGTCTCTGGGAGGATCATCTCTGCATGGTCCAGTCACTCAAAAC 1538

GTGGAGGACAGGAT 77

|||||

GTGGAGGACAGGAT 1555

andard; cDNA; 2839 BP.

AC AAQ54352;  
XX  
DT 25-MAR-2003 (revised)  
DT 23-JUN-1994 (first entry)  
XX  
DE Carcinoembryonic antigen CEA-(b) coding sequence.  
XX  
KW Carcinoembryonic antigen; CEA; tumour marker; colorectal cancer  
KW immunocassay; LoVo cell; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key  
CDS Location/Qualifiers  
FT 5. .2113  
FT /\*tag= a  
FT /product= "CEA-(b)"  
FT /transl\_except= pos:182..184,aa:Glu  
FT /transl\_except= pos:350..352,aa:Asn  
FT /transl\_except= pos:1400..1402,aa:Asn  
FT mat\_peptide  
FT 107..2110  
FT /\*tag= b  
XX  
PN US5274087-A.  
XX  
PD 28-DEC-1993.  
XX  
PF 29-APR-1992; 92US-00876821.  
XX  
PR 13-AUG-1986; 86US-00896361.  
PR 19-FEB-1987; 87US-00016683.  
PR 19-JUN-1987; 87US-00060031.  
PR 12-AUG-1988; 88US-00231741.  
XX  
PA (MOLE-) MOLECULAR DIAGNOSTICS INC.  
XX  
PI Barnett TR, Kamarc ME, Elting JJ;  
XX  
DR WPI; 1994-006741/01.  
DR P-PSDB; AAR54713.  
XX  
PT Nucleic acid encoding carcinoembryonic antigens, - used in det  
PT methods to diagnose different types of cancer.  
XX  
PS Claim 2; Fig 5; 39pp; English.  
XX  
CC CEA-(b) is a member of the carcinoembryonic antigen family. The  
CC sequence AAQ54352 was isolated from a LoVo tumour cell cDNA lib  
CC Plasmid pCllv7 comprising the cDNA insert sequence was deposited  
CC E.coli as ATCC # 67312. (Updated on 25-MAR-2003 to correct PF;  
XX  
SQ Sequence 2839 BP; 790 A; 811 C; 574 G; 664 T; 0 U; 0 Other;  
Query Match 58.4%; Score 45; DB 2; Length 2839;  
Best Local Similarity 74.0%; Pred. No. 4.5e-06;  
Matches 57; Conservative 0; Mismatches 20; Indels 0;  
QY 1 AGCAATCACAGTCTCTGGGAGGATCATCTCTGCATGGTCCAGTCACTCAAACT  
Db 1479 AGCAATCACAGTCTCTGGGAGGATCATCTCTGCATGGTCCAGTCACTCAAACT  
QY 61 CCSTGGAGGACAGGAT 77  
Db 1539 CCSTGGAGGACAGGAT 1555  
RESULT 10  
AAT46062  
ID AAT46062 standard; cDNA; 2839 BP.  
XX  
AC AAT46062;  
XX  
DT 25-MAR-2003 (revised)  
DT 17-MAR-1997 (first entry)

ryonic antigen CEA-b cDNA clone cILV7.

ryonic antigen; CEA-b; colorectal cancer; diagnosis; therapy;

ns.

# Location/Qualifiers

5..2113  
/\*tag= a  
/transl\_except= pos:182..184: aa:Glu  
/transl\_except= pos:350..352: aa:Asn  
/transl\_except= pos:1400..1402: aa:Asn

-A.

96.

33; 93US-00170134.

36; 86US-00896361.

37; 87US-00016683.

37; 87US-00060031.

38; 88US-00231741.

42; 92US-00876821.

AYER CORP.

3, Elting JJ, Kowarsck ME;

-505404/50.

W066872.

ing carcinoembryonic antigen polypeptide - for diagnostic or  
ic use.

3; Fig 5A-G; 32pp; English.

one (AA746062), designated cILV7, encodes carcinoembryonic  
2A-b (AAW06872). It was identified in LoVo cDNA using a  
ed CEA clone LV7 (see also AA746061) fragment probe. A plasmid  
contg. pCLV7 has been deposited as ATCC 67312. CEA gene  
(see also AA746063-64) can be used as probes to isolate other  
the CEA gene family, to determine expression of CEA genes in  
mour types, and to produce recombinant CEA polypeptides, esp.  
useful in immunoassays. (Updated on 25-MAR-2003 to correct PF

839 BP; 790 A; 811 C; 574 G; 664 T; 0 U; 0 Other;

58.4%; Score 45; DB 2; Length 2839;

ilarity 74.0%; Pred. No. 4.5e-06;

Conservative 0; Mismatches 20; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGGGAGGACATCATCTCTGCATGTCAGGTCTAATCTCAAC 60

AGACAATCACAGTCTCTGGGAGGACGTCGACGCTCCATCTCCAGCAACTCAAC 1538

CGTGGAGGACAAAGAT 77

CGTGGAGGACAAAGAT 1555

standard; cDNA; 2459 BP.

01 (first entry)

er associated gene sequence SEQ ID NO:291.

XX Human; cancer associated gene; cancer antigen; detection; can  
KW diagnosis; cytostatic; proliferative; vulnery; immunomodula  
KW antidiabetic; antisthmatic; antirheumatic; antiarthritic; an  
KW antiinflammatory; antithyroid; antiallergic; antibacterial; c  
KW dermatological; neuroprotective; thrombolytic; coagulant; noo  
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; infl  
KW immune disorder; haematopoietic cell disorder; autoimmune dis  
KW allergic reaction; graft versus host disease; organ rejection  
KW haemostatic; thrombolytic; cardiovascular disorder; infection  
KW neurological disease; drug screening; as.

XX Homo sapiens.

OS W0200055350-A1.

XX 21-SEP-2000.

PD 08-MAR-2000; 2000WO-US005882.

PF 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Ruben SM;

PI WPI; 2000-587533/55.

XX P-PSDB; AAB43688.

DR Novel isolated nucleic acids comprising sequences encoding pei  
XX useful for treating or diagnosing e.g. cancer.

XX Claim 1; Page 852; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated prote:  
CC AAB43398 to AAB44239. The proteins can have activities based  
CC tissues and cells the genes are expressed in. Example of acti  
CC include: cytostatic; proliferative; vulnery; immunomodula:  
CC antidiabetic; antisthmatic; antirheumatic; antiarthritic;  
CC antiinflammatory; antithyroid; antiallergic; antibacterial; a  
CC dermatological; neuroprotective; cardiac; thrombolytic; coag  
CC neotropic; vasotropic; antipsoriatic and antiangiogenic. The  
CC polynucleotides and polypeptides can be used for preventing  
CC ameliorating medical conditions and diagnosing pathological c  
CC polynucleotides, polypeptides, antibodies, agonists and antagon  
CC the present invention may be used to treat immune disorders b  
CC or inhibiting the proliferation, differentiation or mobilisat  
CC immune cells, to treat disorders of haematopoietic cells, aut  
CC disorders, allergic reactions, graft versus host disease and  
CC rejection, modulate haemostatic or thrombolytic activity, mod  
CC inflammation, cancers, cardiovascular disorders, neurological  
CC bacterial or viral infections. The peptides, nucleotides, ant  
CC agonists and antagonists may be also be used in drug screens.  
CC AAC78457 and AAB44240 represent sequences used in the exampli  
CC the present invention

XX Sequence 2459 BP; 661 A; 748 C; 517 G; 518 T; 0 U; 15 Other;

Query Match 57.9%; Score 44.6; DB 3; Length 2459;

Best Local Similarity 72.7%; Pred. No. 6.1e-06;

Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACAATCACAGTCTCTGGGAGGACATCATCTCTGCATGTCAGGTCTAATCAAC

DB 1582 AGACAATCACAGTCTCTGGGAGGACGTCGACGCTCCATCTCCAGCAACTCAAC

QY 61 CCGTGGAGGACAAAGAT 77

DB 1642 CCGTGGAGGACAAAGAT 1658

RESULT 12

AAS57592

on tumour protein; colon cancer; gene therapy; cytostatic; ss.

; 2000US-0191597P.  
; 2000US-0202024P.  
; 2000US-0202189P.

IXA CORP.

Xu J, King GE;

11627/70.

tumor proteins and related nucleic acid, useful for treatment,  
diagnosis and monitoring of cancer.

age 232; 299pp; English.

invention relates to the isolation of novel cDNA sequences  
or at least an immunogenic portion of human colon tumor  
proteins. The sequences of the invention are useful in pharmaceu-  
ticals and vaccines for the prevention and treatment of cancers  
of such cancers. They are also useful for the diagnosis and  
monitoring of such cancers. Antibodies to the colon tumor proteins and  
antigen presenting cells that express polynucleotides encoding colon  
tumor proteins can be used to inhibit the development of cancers. T-  
cells react specifically with colon tumor proteins are useful for  
removing tumor cells from samples (e.g. blood) and for cancer treatment.  
The polynucleotide sequences are also useful in gene therapy. AAS57325-  
represent the cDNA sequences of the invention that encode for  
human colon tumour proteins

26 BP; 88 A; 100 C; 78 G; 60 T; 0 U; 0 Other;

Similarity 56.4%; Score 43.4; DB 4; Length 326;

milarity 72.7%; Pred. No. 9.6e-06;

Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACATCACAGTCTCTGGGAGGACATCATCTCTGCATGTCAGTCACTCCAAAC 60  
|||||  
GACATCACAGTCTCTGGGAGGACATCATCTCTGCATGTCAGTCACTCCAAAC 80

CGTGGAGGACAGGAT 77

|||||  
CGTGGAGGACAGGAT 97

standard; cDNA; 326 BP.

2 (first entry)

encoding portion of a human colon tumour protein.

on tumour protein; colon cancer; gene therapy; cytostatic; ss.

ns.

17-R2.

11.

11; 2001WO-US009246.

30; 2000US-0191597P.

00; 2000US-0202024P.

00; 2000US-0202189P.

ORIXA CORP.

J, Xu J, King GE;

-611627/70.

XX New colon tumor proteins and related nucleic acid, useful for t  
PT prevention, diagnosis and monitoring of cancer.

XX Claim 4; Page 275; 299pp; English.

XX Th present invention relates to the isolation of novel cDNA seq  
CC encoding for at least an immunogenic portion of human colon tu  
CC proteins. The sequences of the invention are useful in pharmace  
CC compositions and vaccines for the prevention and treatment of c  
CC such as colon cancer. They are also useful for the diagnosis ar  
CC monitoring of such cancers. Antibodies to the colon tumour prot  
CC antigen presenting cells that express polynucleotides encoding  
CC tumour proteins can be used to inhibit the development of cance  
CC cells that react specifically with colon tumour proteins are us  
CC removing tumour cells from samples (e.g. blood) and for cancer  
CC The polynucleotide sequences are also useful in gene therapy.  
CC AAS58860 represent the cDNA sequences of the invention that en  
CC portions of human colon tumour proteins

XX SQ Sequence 326 BP; 88 A; 100 C; 78 G; 60 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 4; Length 326;  
Best Local Similarity 72.7%; Pred. No. 9.6e-06;  
Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACATCACAGTCTCTGGGAGGACATCATCTCTGCATGTCAGTCACTAAC  
|||||  
DB 21 AGACATCACAGTCTCTGGGAGGACATCATCTCTGCATGTCAGTCACTAAC

QY 61 CGTGGAGGACAGGAT 77

DB 81 CGTGGAGGACAGGAT 97

RESULT 16

AAS58211

ID AAS58211 standard; cDNA; 326 BP.

XX AAS58211;

DT 13-FEB-2002 (first entry)

DE cDNA #887 encoding portion of a human colon tumour protein.

XX Human; colon tumour protein; colon cancer; gene therapy; cytos

XX Homo sapiens.

XX WO200173027-A2.

XX 04-OCT-2001.

XX 22-MAR-2001; 2001WO-US009246.

XX 24-MAR-2000; 2000US-0191597P.

XX 04-MAY-2000; 2000US-0202024P.

XX 05-MAY-2000; 2000US-0202189P.

XX (CORI-) CORIXA CORP.

XX Meagher MJ, Xu J, King GE;

XX WPI; 2001-611627/70.

XX New colon tumor proteins and related nucleic acid, useful for  
PT prevention, diagnosis and monitoring of cancer.

XX Claim 4; Page 197; 299pp; English.

XX Th present invention relates to the isolation of novel cDNA s  
CC encoding for at least an immunogenic portion of human colon t  
CC proteins. The sequences of the invention are useful in pharm

ion relates to a human colon tumour expressed polynucleotide  
ing a polypeptide ((i), AB876791-AB87693) comprising: (i) any  
defined nucleotide sequences (AB87669-AB87689); ((ii)  
s of (i); ((iii) at least 20 contiguous residues of (i); (iv)  
that hybridize to (i), under moderately stringent conditions;  
ces having at least 75% or 90% identity to (i); or (vi)  
variants of (i). The compositions and methods of the present  
are useful for the diagnosis, prevention and/or treatment of  
recticularly colon cancer. (I) can be used in gene therapy and



; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGGGAGCATCATCTCTGCATGTCAGTCTCATTAACCTCAAC 60  
|||||  
AGACAATCACAGTCTCTGGGAGCTGCCCAAGCCTCCATCTCCAGCAACTCAAC 80  
|||||

CCGTGGAGGACAAAGAT 77  
|||||  
CCGTGGAGGACAAAGAT 97  
|||||

standard; cDNA; 572 BP.

32 (first entry)

4 encoding portion of a human colon tumour protein.

lon tumour protein; colon cancer; gene therapy; cytostatic; ss.

ans.

27-A2.

31.

31; 2001WO-US009246.

30; 2000US-0191597P.

30; 2000US-0202024P.

30; 2000US-0202189P.

DRIXA CORP.

J, Xu J, King GE;

-611627/70.

tumor proteins and related nucleic acid, useful for treatment,  
1, diagnosis and monitoring of cancer.

Page 234; 299pp; English.

; invention relates to the isolation of novel cDNA sequences  
for at least an immunogenic portion of human colon tumour  
The sequences of the invention are useful in pharmaceutical  
ons and vaccines for the prevention and treatment of cancers  
olon cancer. They are also useful for the diagnosis and  
of such cancers. Antibodies to the colon tumour proteins and  
representing cells that express polynucleotides encoding colon  
teins can be used to inhibit the development of cancers. T-  
react specifically with colon tumour proteins are useful for  
tumour cells from samples (e.g. blood) and for cancer treatment.  
nucleotides sequences are also useful in gene therapy. AAS57325-  
represent the cDNA sequences of the invention that encode for  
of human colon tumour proteins

372 BP; 141 A; 168 C; 113 G; 125 T; 0 U; 25 Other;

56.4%; Score 43.4; DB 4; Length 572;

ilarity 72.7%; Pred. No. 1.1e-05;

; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGGGAGCATCATCTCTGCATGTCAGTCTCATTAACCTCAAC 60  
|||||  
AGACAATCACAGTCTCTGGGAGCTGCCCAAGCCTCCATCTCCAGCAACTCAAC 80  
|||||

CCGTGGAGGACAAAGAT 77

CCGTGGAGGACAAAGAT 97

RESULT 20  
ABK86206  
ID ABR86206 standard; cDNA; 2019 BP.  
XX AC ABR86206;  
XX DT 24-SEP-2002 (first entry)  
XX DE cDNA encoding human carcino-embryonic antigen (CEA) variant.  
XX KW Human; carcino-embryonic antigen; CEA; cytostatic; antiviral;  
KW immunostimulant; cell-mediated immune response; tumour; breast  
KW virus infection; prostate cancer; colorectal cancer; pancreat  
KW lymphoma; leukaemia; hepatitis; lentivirus; herpesvirus;  
KW human immunodeficiency virus; HIV; flavivirus; pestivirus; ge  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX CDS 7..2019  
FT /\*tag= a  
FT /product= "Carcino-embryonic antigen (CEA) va  
FT the signal sequence deleted"  
XX PN WO200240059-A2.  
XX PD 23-MAY-2002.  
XX PF 01-NOV-2001; 2001WO-US045626.  
XX PR 01-NOV-2000; 2000US-00704232.  
XX PA (AMBI-) AMERICAN FOUND BIOLOGICAL RES INC.  
PA (MNC/) MINCHEFF M S.  
PA (LOUK/) LOUKINOV D I.  
PA (ZOUB/) ZOUBAK S.  
XX FI Mincheff MS, Loukinov DI, Zoubak S;  
XX WPI; 2002-527524/56.  
XX P-PSDB; AAU98922.  
XX PT Inducing a cell-mediated immune response against a target ant-  
PT reducing undesired cells and stimulating presentation of an a  
PT cell, comprises administering a polynucleotide encoding a var  
PT antigen.  
XX PS Disclosure; Page 123-126; 146pp; English.  
XX CC The invention relates to a method of inducing a cell-mediated  
CC response against a cell comprising a target antigen (I) in a ;  
CC treating a subject having undesired cells, for example tumour  
CC virally-infected cells (C), reducing the number of (C) in a g  
CC stimulating presentation of (I) by a cell. This is done by ad  
CC a polynucleotide (II) encoding a variant of (I), so that (II)  
CC in a cell and cell-mediated immune response is induced. The m  
CC used to treat prostate cancer, breast cancer, colorectal cance  
CC pancreatic cancer, as well as lymphomas and leukaemias. The m  
CC also useful in treating chronic viral infections such as those  
CC hepatitisviruses, lentiviruses (including human immunodeficiency  
CC (HIV)), herpesviruses and the flaviviruses and pestiviruses.  
CC sequence represents the coding sequence of human carcino-embry  
CC antigen (CEA) variant, which has the signal sequence deleted,  
CC target antigen in the method of the invention  
XX SQ Sequence 2019 BP; 546 A; 621 C; 425 G; 427 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 6; Length 2019;  
Best Local Similarity 72.7%; Pred. No. 1.1e-05;  
Matches 56; Conservative 0; Mismatches 21; Indels 0;

AGCAATCACAGTCTCTGGGAGCATCATCTCTGCATGTCAGGTCACTAATCCAAAC 60  
 |||||  
 AGCAATCACAGTCTCTGGGAGCTGCCAAGCCCTCCATCTCCAGCAACTCCAAAC 144

CGTGGAGGACAAGGAT 77  
 |||||  
 CGTGGAGGACAAGGAT 1461

standard; DNA; 2031 BP.

D3 (revised)  
 D5 (first entry)

ryonic antigen DNA.

ryonic antigen; CEA; neoplastic diseases; ds.

ms.

Location/Qualifiers  
 le 103..2028  
 /\*tag= a

1.

4.

4; 94EP-00103986.

3; 93EP-00810214.

FFMANN LA ROCHE & CO AG F.

, Pelegriin A, Mach J;

304461/38.

R60619.

ryonic antigen (CEA) derivs - useful as reagents in immunoassay  
 sis of neoplastic diseases.

age 18; 30pp; English.

s the DNA sequence which encodes carcinoembryonic antigen (CEA)  
 CEA is free from cross-reactive CEA-like antigens, it is  
 lly indistinguishable from the solution form of CEA shed from  
 ls, and it is devoid of ethanolamine. AAR60619 can be used in a  
 mposition for detecting neoplastic diseases in biological  
 r in an immunoassay process where it can specifically detect  
 ce of tumour cells in a biological sample e.g. blood. (Updated  
 2003 to correct PN field.)

031 BP; 551 A; 642 C; 417 G; 421 T; 0 U; 0 Other;

ilarity 56.4%; Score 43.4; DB 2; Length 2031;  
 Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GCAATCACAGTCTCTGGGAGCATCATCTCTGCATGTCAGGTCACTAATCCAAAC 60  
 |||||  
 GCAATCACAGTCTCTGGGAGCTGCCAAGCCCTCCATCTCCAGCAACTCCAAAC 1534

CGTGGAGGACAAGGAT 77

CGTGGAGGACAAGGAT 1551

AAT36495  
 ID AAT36495 standard; cDNA; 2059 BP.  
 XX  
 AC AAT36495;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 27-AUG-2003 (revised)  
 DT 13-NOV-1996 (first entry)  
 XX  
 XX Immunogenic carcinoembryonic antigen cDNA.  
 XX  
 DE  
 XX  
 KW Carcinoembryonic antigen; immunogen; breast cancer; lung cano  
 KW colon cancer; therapy; immunotherapy; vaccine; baculovirus; v  
 KW Spodoptera frugiperda; insect; pA9080 ACNPV-CEA; ss.  
 XX  
 XX Autographa californica nucleopolyhedrovirus.  
 OS Homo sapiens.  
 OS Chimeric.  
 XX  
 FH Key Location/Qualifiers  
 FT misc\_feature 1..7  
 FT /\*tag= a  
 FT /function= "7-adenine motif in the polyhedrin  
 FT leader"  
 FT 18..2009  
 FT /\*tag= b  
 FT 18..71  
 FT /\*tag= c  
 FT /function= "ACNPV 61k protein signal sequence"  
 FT mat\_peptide 72..2006  
 FT /\*tag= d  
 FT 72..80  
 FT /\*tag= e  
 FT /function= "3 N-terminal residues of the bacu  
 FT (MGS12) not present in human CEA"  
 FT 81..2006  
 FT /\*tag= f  
 FT /function= "mature CEA"  
 FT primer\_bind complement(81..103)  
 FT /\*tag= g  
 FT /note= "5' primer for CEA amplification"  
 FT primer\_bind 1985..2006  
 FT /\*tag= g  
 FT /note= "3' primer for CEA amplification"  
 FT 2030..2040  
 FT /\*tag= i  
 FT /function= "universal translation termination  
 FT contg. stop codons in all 3 reading frames in  
 FT MGS12"  
 XX  
 XX WO9532286-A2.  
 XX 30-NOV-1995.  
 XX  
 XX 19-MAY-1995; 95WO-US006373.  
 XX  
 XX 20-MAY-1994; 94US-00246981.  
 XX (MICR-) MICROGENESYS INC.  
 XX  
 XX Smith G, Volvovitz F, Hackett C;  
 PI  
 DR WPI; 1996-020581/02.  
 DR P-PSDB; AAR98519.  
 XX  
 XX Immunogenic carcinoembryonic antigen produced using insect cel  
 PT baculovirus expression system - useful in cancer therapy.  
 XX  
 XX Disclosure; Page 47-49; 61pp; English.  
 XX  
 XX A portion (AAT36494) of pA9080 ACNPV-CEA vector codes for reco  
 CC soluble, immunogenic carcinoembryonic antigen (rCEA). (AAR98519  
 CC human CEA gene was cloned from colon adenocarcinoma LS174T (AT

PCR (see also AAT36493-94) and modified for expression in insect replacement of the signal sequence with a 61 kDa baculovirus sequence, and deletion of DNA encoding the hydrophobic C-terminal. The construct was inserted into baculovirus vector pMGs12 contg. Irlin promoter, giving plasmid pA9080 AcNPV-CEA. rCEA was produced in Sf900+ insect cells grown in serum-free media and purified to a purity of over 95%. It can be used as an immunogen in protect against breast, lung or colon cancer. (Updated on 27- to correct OS field.) (Updated on 16-Oct-2003 to standardise OS

2059 BP; 572 A; 623 C; 425 G; 439 T; 0 U; 0 Other;

Similarity 56.4%; Score 43.4; DB 2; Length 2059;  
; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

AGACAATCAGTCTCTGGGAGCAGCATCTCTGCGAGCTGCTCCATCTCCAGCAGCATCTCCAAAC 1512  
AGACAATCAGTCTCTGGGAGCAGCATCTCTGCGAGCTGCTCCATCTCCAGCAGCATCTCCAAAC 1512

CGGTGGAGGACAGGAT 77

|||||  
CGGTGGAGGACAGGAT 1529

standard; cDNA; 2097 BP.

96 (first entry)

oryonic antigen cDNA.

mplify; polymerase chain reaction; PCR; human;  
lycoprotein; BGP; membrane-bound; carcinoembryonic antigen; CEA;  
protein; P1A3 epitope; anti-P1A3 antibody;  
l carcinoma; monoclonal antibody; ss.

ans.

Location/Qualifiers

le 1..90

/\*tag= a

le 91..2094

/\*tag= b

-Al.

15.

14; 94WO-GB001816.

13; 93GB-00017423.

IPERIAL CANCER RES TECHNOLOGY.

Durbin H, Snary D, Stewart LM, Young S, Bates PA;

106813/14.

WR65168.

les which bind carcinoembryonic antigen - used for the  
and treatment of colorectal carcinoma and for isolation and  
ons.

; Fig 12; 67pp; English.

ces given in AA082805-07 encode human biliary glycoprotein  
on-specific cross-reacting antigen (NCA) and human membrane-  
inoembryonic antigen (CEA) respectively. BGP and CAE may be

CC used in the preparation of chimaeric proteins which may be used  
CC identify the P1A3 antigen and to raise the antibody of the ;  
CC The preferred chimaeras are given in AAR77435-38. NCA may be  
CC identify regions of importance in the CEA protein by compari  
CC on the amino acid side chains with that in the binding regio  
CC P1A3 antibody (see also AAQ82785-90). The P1A3 epitope was  
CC an epitope of CEA within the B3-GPI region. Chimaeras in whic  
CC was not membrane bound did not react with anti-P1A3 antibod  
CC indicating that the P1A3 epitope is not present in non-mem  
CC hybrid proteins. Antibodies which recognise the P1A3 epitope  
CC the detection of well and poorly differentiated colorectal ce  
CC The isolation of the specific P1A3 epitope allows the develo  
CC monoclonal antibodies specific for colorectal carcinoma. They  
CC in the study, isolation and purification of molecules to whic  
CC specifically bind and the imaging and treatment of cells exhi  
CC molecules

XX Sequence 2097 BP; 559 A; 656 C; 439 G; 443 T; 0 U; 0 Other;  
SQ

Query Match 56.4%; Score 43.4; DB 2; Length 2097;  
Best Local Similarity 72.7%; Pred. No. 1.6e-05;  
Matches 56; Conservative 0; Mismatches 21; Indels 0

QY 1 AGACAATCAGTCTCTGGGAGCAGCATCTCTGCGAGCTGCTCCATCTCCAGGTCATAA  
Db 1463 AGACAATCAGTCTCTGGGAGCTGCCAAGCCCTCCATCTCCAGCAACA

QY 61 CGGTGGAGGACAGGAT 77

Db 1523 CGGTGGAGGACAGGAT 1539

RESULT 24

AAI72497

ID AAI72497 standard; DNA; 2105 BP.

AC AAI72497;

DT 16-MAY-2002 (first entry)

XX CEA agonist coding sequence #2.

XX CAP-1; epitope; carcinoembryonic antigen; CEA; agonist; immun  
KW carcinoma; gastrointestinal; breast; pancreatic; bladder; ova  
KW prostate; cancer; gene; ss.

OS Synthetic.

XX

Key Location/Qualifiers

FT CDS 1..2105

/\*tag= a

/product= "CEA agonist polypeptide"

PN WO200210379-A2.

PD 07-FEB-2002.

XX 27-JUL-2001; 2001WO-CA001092.

XX 31-JUL-2000; 2000US-0222043P.

XX (AVET ) AVENTIS PASTEUR LTD.  
XX (THER-) THERION BIOLOGICS.

XX (USSH ) US NAT CANCER INST.

XX Berinstein N, Tartaglia J, Tine JA, Panicali DL, Gritz L;  
XX Schlom J;

XX WPI; 2002-206189/26.

XX Carcinoembryonic antigen agonist polypeptide for inducing an i  
PT response in animal against antigen and for inhibiting an epit  
PT expressing carcinoma cell, comprises a modified antigen epitop

Page 66-67; 69pp; English.

ence encodes the carcinoembryonic antigen (CEA) agonist le of the invention. This sequence represents the sequence given (ID listing in the specification, and does not directly encode a polypeptide given in AAB47919. The CEA agonist contains a AP-1 epitope of CEA, in which position 6 is modified from Asp to increase its immunogenicity. The CEA agonist polypeptide of the or DNA encoding it, are useful for: (i) inducing an immune n an animal directed against a CEA protein or fragment, CEA CEA epitope, a modified CEA epitope, cells expressing or CEA protein or fragment; and (ii) inhibiting a CEA epitope carcinoma cell, which is a gastrointestinal, breast, bladder, ovarian, lung or prostate carcinoma cell in a ence is useful for manufacture of a medicament for the of cancer

105 BP; 555 A; 658 C; 441 G; 451 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 6; Length 2105;  
Best Local Similarity 72.7%; Pred. No. 1.6e-05;

Conservative 0; Mismatches 21; Indels 0; Gaps 0;

3ACATATCAGTCTCTGCGAGGATCATCTCTGATGTCAGGTCTCAATCTCAAAAC 60  
3ACATATCAGTCTCTGCGAGGATCGCCCAAGCCCTCCATCTCCAGCACACTCCAAAC 1530

3GTGGAGGACAAGGAT 77

1532 CCGTGGAGGACAAGGAT 1547

Standard; cDNA; 2106 BP.

(first entry)

carcinoembryonic antigen (CEA) encoding cDNA SEQ ID NO:111.

ovirus; poxvirus; alphavirus; immune response; gp100;  
gen; CEA; carcinoembryonic antigen; immunostimulant;  
immunotherapy; interferon-gamma; IFN-gamma; Cancer; ss.

d.

-Al.

2000WO-CA001253.

93US-0160879P.

2000US-022325P.

NTIS PASTEUR LTD.

N, Tartaglia J, Moingeon P, Barber B;

08587/32.

97817.

immune response to tumor antigen, useful in immunotherapy of administering the antigen to a lymphatic site.

Fig 8; 60pp; English.

invention describes a method for inducing an immune response, to a tumour antigen (Ag) comprising administering Ag, or d (I) that encodes it, to a lymphatic site. Cynomolgus monkeys

CC (Macaca fascicularis) were injected with a modified form of gp  
CC (a) into the left inguinal lymph node or (b) subcutaneously,  
CC of (a) developed a cell-mediated response (indicated by produ  
CC interferon-gamma from T lymphocytes when exposed to gp100 pep  
CC only 2 of 4 animals of (b) did so. Also animals in (a) produ  
CC greater antibody response to gp100. The method is used in imm  
CC of a wide range of cancers through induction of a specific im  
CC response (humoral and cellular) against the tumour antigens. I  
CC administered to a lymphatic site, Ag (or (I)) induces a strong  
CC response than administration by other routes and may also bre  
CC to Ag. AAB97708 and AAB97709 represent gp100 epitopes; AAB977:  
CC AAB97815 represent peptides derived from gp100 which stimulat  
CC (IFN)-gamma production; AAB20120 encodes the modified gp100 p  
CC in AAB97816; AAB20121 encodes the modified carcinoembryonic ar  
CC protein given in AAB97817; and AAB97818 represents a CEA modif  
CC peptide, all of which are used in the exemplification of the f  
CC invention  
XX

SQ Sequence 2106 BP; 559 A; 659 C; 442 G; 446 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 4; Length 2106;  
Best Local Similarity 72.7%; Pred. No. 1.6e-05;  
Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACATATCAGTCTCTGCGAGGATCATCTCTGATGTCAGGTCTCAATCTCAAAAC  
Db 1472 AGACATATCAGTCTCTGCGAGGATCGCCCAAGCCCTCCATCTCCAGCACACTCCAAAC

QY 61 CCGTGGAGGACAAGGAT 77

Db 1532 CCGTGGAGGACAAGGAT 1548

RESULT 26

AAD07347

ID AAD07347 standard; DNA; 2106 BP.

XX AAD07347;

XX 18-SEP-2001 (first entry)

XX Modified carcinoembryonic antigen (CEA) DNA.

XX Tumour-associated antigen; TAA; cytostatic; vaccine; gene ther  
KW immune response; tetanus toxoid; TT; diphtheria toxoid; DT; pr  
KW cancer; therapeutic; carcinoembryonic antigen; CEA; ds.  
XX Synthetic.

OS

XX Key

XX CDS Location/Qualifiers

FT 1..2106

FT /\*tag= a

FT /product= "Modified carcinoembryonic antigen (C

XX WO200149317-A2.

XX 12-JUL-2001.

XX 05-JAN-2001; 2001WO-CA000005.

XX 05-JAN-2000; 2000US-0174587P.

XX (AVET ) AVENTIS PASTEUR LTD.

XX Entage P, Barber BH, Sambhara S, Sia CDY;

XX WPI; 2001-441790/47.

XX P-PSDB; AAE05117.

XX Enhancing immune response to antigen such as tumor antigen for  
PT cancer in an animal involves administering an inducing agent to  
PT animal followed by administering inducing agent-antigen mixture  
XX

ig 3; 62pp; English.

ion relates to a method of enhancing an immune response against oiated antigens (TAAs), such as GP100 and carcinoembryonic EA) in an animal. The method involves priming of the animal ducing agent such as tetanus toxoid (TT) or diphtheria toxoid equently followed by administration of an inducing agent-xture. The method provides the enhancement or augmentation of response to the antigen and/or improves a vaccination protocol g use of less antigen. The immunisation of the animal with oiated antigen is useful for the prophylactic or therapeutic of cancer. The present DNA sequence encodes modified ryonic antigen (CEA) related to the invention

106 BP; 559 A; 659 C; 442 G; 446 T; 0 U; 0 Other;

milarity 56.4%; Score 43.4; DB 5; Length 2106;

Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAAATCACAGTCTCTGCGGAGCATCATCTCTGCGATGTCAGGTCTATACTCCAAAC 60  
|||||  
GACAAATCACAGTCTCTGCGGAGTGGCCAGCCCTCCATCTCCAGCAACAATCCAAAC 1531

CGTGGAGGACAGGAT 77

|||||

CGTGGAGGACAGGAT 1548

standard; DNA; 2106 BP.

2 (first entry)

it coding sequence #1.

tope; carcinoembryonic antigen; CEA; agonist; immune response;  
gastrointestinal; breast; pancreatic; bladder; ovarian; lung;  
cancer; gene; ds.

Location/Qualifiers

1. 2106

/\*tag= a

/product= "CEA agonist polypeptide"

'9-A2.

12.

11; 2001WO-CA001092.

10; 2000US-0222043P.

VENTIS PASTEUR LTD.

HERION BIOLOGICS.

; NAT CANCER INST.

1 N, Tartaglia J, Tine JA, Panicali DL, Gritz L;

-206189/26.

1B47918.

ryonic antigen agonist polypeptide for inducing an immune  
in animal against antigen and for inhibiting an epitope antigen  
; carcinoma cell, comprises a modified antigen epitope.

Fig 1; 69pp; English.

XX

CC This sequence encodes the carcinoembryonic antigen (CEA) agoni  
CC polypeptide of the invention. This sequence represents the seq  
CC in the figures in the specification, and it directly encodes t  
CC agonist polypeptide given in ABA47918. The CEA agonist contain  
CC modified CAP-1 epitope of CEA, in which position 6 is modified  
CC to Asn to increase its immunogenicity. The CEA agonist polypep  
CC invention, or DNA encoding it, are useful for: (i) inducing an  
CC response in an animal directed against a CEA protein or fragme  
CC agonist, a CEA epitope, a modified CEA epitope, cells expressi  
CC binding a CEA protein or fragment; and (ii) inhibiting a CEA e  
CC expressing carcinoma cell, which is a gastrointestinal, breast  
CC pancreatic, bladder, ovarian, lung or prostate carcinoma cell  
CC patient, hence is useful for manufacture of a medicament for t  
CC treatment of cancer

XX SQ Sequence 2106 BP; 559 A; 658 C; 442 G; 447 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 6; Length 2106;

Best Local Similarity 72.7%; Pred. No. 1.6e-05;

Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGCAATCACAGTCTCTGCGGAGCATCATCTCTGCGATGTCAGGTCTATAAC  
|||||  
Db 1472 AGCAATCACAGTCTCTGCGGAGTGGCCAGCCCTCCATCTCCAGCAAC

QY 61 CGTGGAGGACAGGAT 77

|||||

Db 1532 CGTGGAGGACAGGAT 1548

RESULT 28

ADE13860

ID ADE13860 standard; DNA; 2106 BP.

XX

AC ADE13860;

XX

DT 29-JAN-2004 (first entry)

XX

DE CEA-CAP6D nucleotide sequence SEQ ID NO:23.

XX

KW carcinoembryonic antigen; CEA; CEA(6D)-1; 2; cytostatic; vacci

XX

KW tumour antigen; immunotherapy; gene; ds.

XX

OS Unidentified.

XX

PN WO2003085087-A2.

XX

PD 16-OCT-2003.

XX

PF 09-APR-2003; 2003WO-US010916.

XX

PR 09-APR-2002; 2002US-0372972P.

XX

PA (AVET ) AVENTIS PASTEUR LTD.

XX

PA (THER-) THERION BIOLOGICS INC.

XX

PI Parrington M, Zhang L, Rovinski B, Gritz LR, Greenhalgh T;

XX

PT WPI; 2003-877029/81.

XX

PT New isolated DNA molecule comprising the carcinoembryonic anti

XX

PT 1,2 sequence, useful for diagnosing, preventing and treating c

XX

PT determining the effectiveness of a chemotherapeutic or other t

XX

PS Example 1; SEQ ID NO 23; 56pp; English.

XX

CC The present invention describes an isolated DNA molecule compri

XX

CC carcinoembryonic antigen (CEA) (6D)-1,2 sequence of 2106 bp (

XX

CC ADE13861), or its fragment. Also described: (1) an expression

XX

CC comprising the nucleic acid sequence CEA(6D)-1,2, or its fragm

XX

CC describes above; (2) a composition comprising the expression

pharmaceutical carrier; and (3) preventing or treating cancer by administering to a host the expression vector of (1). CEA(6D)-toxic activity, and can be used in vaccines. The CEA(6D)-1,2 and target polypeptide are useful for diagnosing, preventing cancer, predicting prognosis, or determining the effectiveness of a chemotherapeutic or other treatment regimen. The vector may be used for the insertion and expression of CEA(6D) in a cell encoding tumour antigens for the immunotherapeutic use of cancer. The target polypeptides are useful in generating used in screening assays or for immunotherapy. The present represents the CEA-CAP6D nucleotide sequence, which is given in with CEA(6D)-1,2 in the exemplification of the present

106 BP; 559 A; 659 C; 442 G; 446 T; 0 U; 0 Other;

milarity 56.4%; Score 43.4; DB 9; Length 2106;  
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGGGAGCATCATCCTCTGCATGGTCAAGTCATTAATCCAAAC 60  
|||||  
GACAATCACAGTCTCTGGGAGCATCATCCTCTGCATGGTCAAGTCATTAATCCAAAC 1531  
CTGGGAGGACAAGGAT 77  
CTGGGAGGACAAGGAT 1548

standard; cDNA; 2109 BP.

2 (first entry)

cyonic antigen coding sequence.

cyonic antigen; CEA; epitope; human leucocyte antigen; HLA; compatibility complex; MHC; CD4; T lymphocyte; cancer; tumour; antitumour; therapy; diagnosis; human; gene; ss.

15.

Location/Qualifiers  
1..2109  
/\*tag= a  
/product= "CEA"

1-A2.

1.

2001WO-US028467.

2000US-0232185P.

OS INC.

omlinson A;

62345/39.

175931.

human leucocyte antigen (HLA) class II-restricted epitopes, isolating a major histocompatibility complex (MHC) molecule peptide and eluting the peptide.

Fig 1A; 98pp; English.

sequence is the coding sequence for human carcinoembryonic A). Epstein-Barr virus-transformed B cell lines expressing CEA

CC encoding cDNA were used to identify human leucocyte antigen (II binding CEA peptide epitopes. The invention provides methods identifying peptide epitopes that activate CD4+ T lymphocyte involved in the initiation, promotion, or exacerbation of certain diseases. It is based on the discovery that it is possible to human leucocyte antigen class II restricted epitopes naturally, by antigen-presenting cells transfected with DNA encoding a peptide which the epitopes are derived. The method has 2 phases: immun mass fingerprinting, and epitope verification. By applying the the tumour antigen CEA, CEA-derived peptides (see ABE75917-26) identified as epitopes that could be involved in the pathogenesis cancer in human patients expressing the DR4 or DR1 class II major histocompatibility complex (MHC) allele. The epitopes, and alternative peptide ligands (APL), are used in claimed methods of activating reactivity, altering T cell response, diagnosis, and identifying reagent for diagnosing cancer. Also claimed are methods of cancer therapy, and of identifying a class II MHC-binding fragment of antigen. APLs are also useful as vaccines

SQ Sequence 2109 BP; 561 A; 659 C; 442 G; 447 T; 0 U; 0 Other;  
Query Match 56.4%; Score 43.4; DB 6; Length 2109;  
Best Local Similarity 72.7%; Pred. No. 1.6e-05;  
Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACAATCACAGTCTCTGGGAGCATCATCCTCTGCATGGTCAAGTCATTAAC  
|||||  
Db 1475 AGACAATCACAGTCTCTGGGAGCATCATCCTCTGCATGGTCAAGTCATTAAC  
|||||  
QY 61 CGGTGGAGGACAAGGAT 77  
|||||  
Db 1535 CGGTGGAGGACAAGGAT 1551

RESULT 30

AAT33302

ID AAT33302 standard; cDNA; 2220 BP.

XX

AC AAT33302;

XX

DT 12-NOV-1996 (first entry)

XX

DE Carcinoembryonic antigen gene.

XX

KW Poliovirus; encapsidation; capsid; vaccine; genetic immunisation; carcinoembryonic antigen; CEA; tumour; cancer therapy; ss.

XX

OS Homo sapiens.

XX

FT Key Location/Qualifiers  
CDS 1..2206  
FT /\*tag= a

XX W09625173-A1.

XX

PD 22-AUG-1996.

XX

PF 13-FEB-1996; 96WO-US001895.

XX

PR 15-FEB-1995; 95US-00389459.

XX

PA (UABR-) UAB RES FOUND.

XX

PI Morrow CD, Porter DC, Ansardi DC;

XX

DR WPI; 1996-393136/39.

XX

DR P-FSDB; AAW00182.

XX

PT Encapsidation of recombinant polio-virus nucleic acid for use in - using a polio-virus nucleic acid which lacks the P1 capsid re

XX

PS an expression system which provides the region.

PS Example 7; Page 68-71; 108pp; English.

quence (AAT33302) codes for the human carcinoembryonic antigen (CEA). The CEA can be used to substitute the P1 capsid gene of the recombinant poliovirus nucleic acids (rPNAs). Such rPNAs are then introduced into a host cell together with a poliovirus or plasmid vector encoding the poliovirus P1 capsid protein. Encapsidated rPNAs are useful for genetic immunisation, stimulating an immune response to CEA as a means of cancer therapy.

Sequence 2349 BP; 640 A; 693 C; 477 G; 539 T; 0 U; 0 Other;  
 Similarity 56.4%; Score 43.4; DB 2; Length 2349;  
 Mismatches 0; Conservative 0; Indels 0; Gaps 0;  
 GACAAATCACAGTCTCTCGGAGGCTGCCAAGCCCTCCATCTCCAGCACACTCCAAAC 1630  
 GACAAATCACAGTCTCTCGGAGGCTGCCAAGCCCTCCATCTCCAGCACACTCCAAAC 1630  
 CGTGGAGGACAGGAT 77  
 CGTGGAGGACAGGAT 1647

standard; DNA; 2349 BP.

13 (revised)  
 15 (first entry)

expression cassette from pH6.CEA.HA.

chain reaction; primer; amplify; NVVAC; ALVAC; recombinant; plasmid; pCEM.CEA; vaccinia; H6 promoter; amplify; primer; transcription termination signal; TSNT; FC33 strain; ALVAC C3; 42K entomopox promoter; VQCP3L; donor plasmid; pSD414; polymerase chain reaction; poxvirus; antigenic response; cal response; pathogen; ss.

Location/Qualifiers  
 1..59  
 /tag= a  
 /note= "Flanking sequence"  
 60..183  
 /tag= b  
 /note= "Vaccinia H6 promoter"  
 184..2292  
 /tag= c  
 /product= "CEA"  
 2293..2349  
 /tag= d  
 /note= "Flanking sequence"

-A1.

14.

14; 94WO-US000088.

13; 93US-00007115.

14; 94US-00184009.

ROGENETICS CORP.

3; Tartaglia J, Cox WI;

-263767/32.

XX Attenuated recombinant virus used for cancer therapy - compris  
 encoding cytokine and/or tumour associated antigen.  
 XX Example 17; Fig 23; 232pp; English.  
 XX The sequences given in AA067868-69 represent H6/CEA expression  
 and flanking regions from plasmids pH6.CEA.C3.2 and pH6.CEA.H6  
 respectively. These sequences were used in the construction of  
 ALVAC-based recombinant viruses containing the CEA gene. The  
 5' and 3' untranslated regions (UTRs) are contained in plasmid  
 The 5' end of the CEA construct was modified to remove the 5'  
 place the vaccinia H6 promoter before the ATG initiation codon  
 primers CEA1 and CEA2. The resulting fragment links the 3' 30  
 H6 promoter to the CEA initiation codon and the next 28 bp of  
 gene. The 3' UTR was also removed and replaced by the vaccinia  
 transcription termination signal (TSNT) followed by a series of  
 restriction sites, using primers CEA3 and CEA4. The SERA cDNA  
 isolated from the FC33 strain of P. faiciparum and modified to  
 the TSNT sequence and a series of restriction sites, at the 3'  
 5' end was modified by replacing the 5' untranslated region (U  
 several restriction sites and the 42K entomopox promoter. This  
 expression cassette was cloned into a BamHI/XhoI-digested VQCI  
 fragment to generate an ALVAC C3 donor plasmid. The pSD553 vac  
 plasmid was then constructed by PCR using pSD414 as a template  
 may also be used in the production of the ALVAC C3 donor plas  
 gene sequence was introduced into ALVAC by recombination betw  
 linearised CEA.C3 donor plasmid and ALVAC rescuing plasmid. The  
 viruses may be used in a composition for inducing an antigenic  
 immunological response, ie. for immunisation against pathogens  
 CC on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 2349 BP; 640 A; 693 C; 477 G; 539 T; 0 U; 0 Other;  
 Query Match 56.4%; Score 43.4; DB 2; Length 2349;  
 Best Local Similarity 72.7%; Pred. No. 1.6e-05;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0;  
 Qy 1 AGCAATCACAGTCTCTCGGAGGCTGCCAAGCCCTCCATCTCCAGCACAA  
 Db 1658 AGCAATCACAGTCTCTCGGAGGCTGCCAAGCCCTCCATCTCCAGCACAA  
 Qy 61 CCGTGGAGGACAGGAT 77  
 Db 1718 CCGTGGAGGACAGGAT 1734  
 RESULT 32  
 AAZ08470  
 ID AAZ08470 standard; DNA; 2349 BP.  
 XX  
 AC AAZ08470;  
 XX  
 DT 19-OCT-1999 (first entry)  
 DE  
 DE H6/CEA expression cassette and flanking regions from pH6.CEA.F  
 XX Attenuated recombinant virus; cytokine; tumour associated anti  
 NVVAC recombinant virus; ALVAC recombinant virus; gene therapy  
 cancer; tumour necrosis factor; nuclear phosphoprotein; p53; I  
 interleukin; interferon; IFN-gamma; IL-4; melanoma associated  
 carcinoembryonic antigen; immunisation; antigenic; poxvirus; i  
 immunological response; immunotherapy; vaccine; Newcastle Dis  
 XX  
 OS Synthetic.  
 OS Vaccinia virus.  
 XX  
 PN US5942235-A.  
 XX  
 PD 24-AUG-1999.  
 XX  
 PF 02-JUN-1995; 95US-00458356.  
 XX

11; 81US-00334456.  
12; 82US-00446824.  
14; 84US-00622135.  
17; 87US-00090209.  
17; 87US-00090711.  
17; 87US-00110335.  
18; 88US-00186054.  
18; 88US-00234390.  
10; 90US-00537882.  
10; 90US-00537890.  
11; 91US-00805567.  
12; 92US-00847977.  
12; 92US-00847951.  
2; 92US-00881995.  
2; 92US-00918278.  
3; 93US-00007115.  
4; 94US-00184009.  
4; 94US-00228926.  
4; 94US-00306259.

ALTH RES INC.

;

493494/41.

t poxviruses comprising exogenous DNA encoding antigenic ts useful in immunotherapy to immunize against cancers and ases such as influenza, Newcastle Disease and rabies.

; Fig 23; 163pp; English.

t invention describes a recombinant poxvirus (I), comprising DNA encoding an antigenic determinant of a pathogen which is used in vivo in infected host cells after administration to a d therefore induces an immunological response. (I) may be used te patients against a wide range of diseases and disorders on the type of antigen encoded by the exogenous DNA. (I) may be vaccine against diseases such as rabies, influenza and disease. It is particularly useful for immunising against ne poxvirus (I) also provides a means of manipulating s and tumour cells for use in cell-based immunotherapeutic ad viruses. (I) also have enhanced safety compared to recombinant poxvirus vaccines. This increased level of safety e possibility of a 'runaway' infection in the host and reduces of transmission from vaccinated to unvaccinated individuals nation of the environment. The present sequence represents a resion cassette and flanking regions from pH6.CEA.HA used in ification of the present invention

149 BP; 640 A; 693 C; 477 G; 539 T; 0 U; 0 Other;

56.4%; Score 43.4; DB 2; Length 2349;  
ilarity 72.7%; Pred. No. 1.6e-05;  
Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
ACAAATCAGTCTCTCGGAGCATCATCTCTGCGATGTCAGTCTAATCTCAAC 60  
ACAAATCAGTCTCTCGGAGCATCTCCCAAGCCCTCATCTCCAGCAACAATCAAC 1717  
XGTGGAGGACAAGGAT 77  
XGTGGAGGACAAGGAT 1734

andard; DNA; 2389 BP.

(first entry)

XX Human carcino-embryonal antigen encoding DNA.  
DE Human; MUC1; carcino-embryonal antigen; CEA; treatment; tumor  
XX tumor antigen; anticancer; vaccine; immunotherapy; ds.  
KW Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
CDS 1..2389  
FT /\*tag= a  
FT /product= "carcino-embryonal antigen"  
FT /transl\_exception= (pos:847..849,aa:Asp)  
FT /note= "This sequence contains in-frame stop  
XX W0200124832-A2.  
XX  
XX  
PD 12-APR-2001.  
XX  
XX 26-SEP-2000; 2000WO-DE003443.  
XX  
XX 27-SEP-1999; 99DE-01048105.  
XX  
XX (PECH/) PECHER G.  
XX  
XX Pecher G;  
PI  
XX  
XX WPI; 2001-266243/27.  
XX P-PSDB; AAB70880.  
XX  
XX Composition for treating or preventing tumors, comprises plasr  
XX optionally recombinant viruses, that express mucin and carcino  
XX antigen.  
XX  
XX Claim 1; Fig 2; 11pp; German.  
XX  
XX This invention describes a novel composition (A) for treating  
XX preventing human tumors that express the tumor antigens mucin  
XX carcino-embryonal antigen (CEA). The composition has anticance  
XX comprises (1) a plasmid containing the human mucin gene MUC1 (  
XX active fragments or a sequence that expresses at least three r  
XX specified sequence (1)); and/or (2) a plasmid containing the g  
XX human CEA. (A) optionally used in combination with recombinant  
XX that expresses the same tumor antigens, are useful as vaccines f  
XX immunotherapy and prevention of tumors that express mucin and/  
XX combination of antigens and vectors improves activity of the c  
XX The method does not require manipulation of tumor cells and is  
XX specific for a particular patient. The vaccines can be adminis  
XX directly in vivo. This sequence encodes the human carcino-emb  
XX antigen (CEA) protein used in the method of the invention  
XX  
SQ Sequence 2389 BP; 637 A; 730 C; 506 G; 516 T; 0 U; 0 Other;  
Query Match 56.4%; Score 43.4; DB 4; Length 2389;  
Best Local Similarity 72.7%; Pred. No. 1.6e-05;  
Matches 56; Conservative 0; Mismatches 21; Indels 0;  
QY 1 AGACAATCAGTCTCTCGGAGCATCATCTCTGCGATGTCAGTCTAATCAAC  
Db 1475 AGACAATCAGTCTCTCGGAGCATCTCCCAAGCCCTCATCTCCAGCAACAAC  
QY 61 CCGTGAGGACAAGGAT 77  
Db 1535 CCGTGAGGACAAGGAT 1551  
RESULT 34  
AAQ67868  
ID AAQ67868 standard; DNA; 2434 BP.  
XX  
XX AAQ67868;  
XX  
XX 25-MAR-2003 (revised)



```

95 (first entry)
pression cassette from pH6.CEA.C3.2.
e chain reaction; primer; amplify; NYVAC; ALVAC; recombinant;
A; plasmid; pGEM.CEA; vaccinia; H6 promoter; amplify; primer;
scription termination signal; TSNT; FC33 strain; ALVAC C3;
arum; 42K entomopox promoter; VQCP3L; donor plasmid; PSD414;
polymerase chain reaction; poxvirus; antigenic response;
ical response; pathogen; ss.

.
Location/Qualifiers
ure 1..56
/*tag= a
/note= "Flanking sequence"
57..180
/*tag= b
/note= "Vaccinia H6 promoter"
181..2289
/*tag= c
/product= "CEA"
2290..2434
/*tag= d
/note= "Flanking sequence"

ure
-AL.
94.
94; 94WO-US000888.
93; 93US-00007115.
94; 94US-00184009.
IROGENETICS CORP.
3; Tartaglia J, Cox WI;
-263767/32.
i recombinant virus used for cancer therapy - comprises DNA
cytokine and/or tumour associated antigen.
7; Fig 22; 232pp; English.
ices given in AAQ57869-69 represent H6/CEA expression cassettes
ing regions from plasmids pH6.CEA.C3.2 and pH6.CEA.HA
ly. These sequences were used in the construction of NYVAC- and
d recombinant viruses containing the CEA gene. The CEA gene and
ntranslated regions (UTRs) are contained in plasmid pGEM.CEA.
l of the CEA construct was modified to remove the 5' UTR and
vaccinia H6 promoter before the ATG initiation codon using
AI and CE42. The resulting fragment links the 3' 30 bp of the
r to the CEA initiation codon and the next 28 bp of the CEA
3' UTR was also removed and replaced by the vaccinia early
ion termination signal (TSNT) followed by a series of
n sites, using primers CE43 and CE44. The SERA cDNA was
from the FC33 strain of P. falciparum and modified to include
sequence and a series of restriction sites, at the 3' end. The
i modified by replacing the 5' untranslated region (UTR) with
striction sites and the 42K entomopox promoter. This SERA
l cassette was cloned into a BamHI/XhoI-digested VQCP3L vector
o generate an ALVAC C3 donor plasmid. The PSD553 vaccinia donor
us then constructed by PCR using PSD414 as a template. PSD553
be used in the production of the ALVAC C3 donor plasmid. The CEA
ence was introduced into ALVAC by recombination between NotI-
i CEA.C3 donor plasmid and ALVAC rescuing plasmid. The resulting
y be used in a composition for inducing an antigenic or
cal response, ie. for immunisation against pathogens. (Updated
2003 to correct PN field.)
SQ Sequence 2434 BP; 668 A; 704 C; 495 G; 567 T; 0 U; 0 Other;
Query Match 56.4%; Score 43.4; DB 2; Length 2434;
Best Local Similarity 72.7%; Pred. No. 1.7e-05;
Matches 56; Conservative 0; Mismatches 21; Indels (
QY 1 AGACAATCACAGTCTCTGGGAAGCATCATCTCTGTCATGTCAGGTGTCATAT
Db 1655 AGACAATCACAGTCTCTGGGAAGCATCATCTCTGTCATGTCAGGTGTCATAT
QY 61 CCGTGAGGACCAAGGAT 77
Db 1715 CCGTGAGGACCAAGGAT 1731
RESULT 35
ARZ08469
ID AAZ08469 standard; DNA; 2434 BP.
XX AC AAZ08469;
XX 19-OCT-1999 (first entry)
XX DE H6/CEA expression cassette and flanking regions from pH6.CEA.
KW Attenuated recombinant virus; cytokine; tumour associated ant
KW NYVAC recombinant virus; ALVAC recombinant virus; gene therap
KW cancer; tumour necrosis factor; nuclear phosphoprotein; p53;
KW interleukin; interferon; IFN-gamma; IL-4; melanoma associated
KW carcinoembryonic antigen; immunisation; antigenic; poxvirus;
KW immunological response; immunotherapy; vaccine; Newcastle Dis
XX OS Synthetic.
XX OS Vaccinia virus.
XX US942235-A.
XX PD 24-AUG-1999.
XX PF 02-JUN-1995; 95US-00458356.
XX PR 24-DEC-1981; 81US-00334456.
XX PR 08-DEC-1982; 82US-00446824.
XX PR 19-JUN-1984; 84US-00622135.
XX PR 27-AUG-1987; 87US-00090209.
XX PR 28-AUG-1987; 87US-00090711.
XX PR 20-OCT-1987; 87US-00110335.
XX PR 25-APR-1988; 88US-00186054.
XX PR 23-AUG-1988; 88US-00234390.
XX PR 14-JUN-1990; 90US-00537882.
XX PR 16-DEC-1991; 91US-00805567.
XX PR 03-MAR-1992; 92US-00847977.
XX PR 04-MAY-1992; 92US-00847951.
XX PR 22-JUL-1992; 92US-00881995.
XX PR 20-JAN-1993; 93US-00007115.
XX PR 19-JAN-1994; 94US-00184009.
XX PR 14-APR-1994; 94US-00228926.
XX PR 13-SEP-1994; 94US-00306259.
XX (HEAL-) HEALTH RES INC.
XX PI Paoletti E;
XX WPI; 1999-493494/41.
XX Recombinant poxviruses comprising exogenous DNA encoding antic
XX determinants useful in immunotherapy to immunize against cance
XX other diseases such as influenza, Newcastle Disease and rabies
XX Example 17; Fig 22; 163pp; English.

```

A (SUNR) SUNTORY LTD.

245625/35.  
P95424.

inoembryonic antigen - used to produce antibodies and detect sue without reacting to CEA-related antibody.

; Page ?; 15pp; Japanese.

carcinoembryonic antigen (CEA) gene is used to express CEA. raised against human CEA are used to detect tumor tissue cinoma). See also AAP95424

928 BP; 824 A; 824 C; 584 G; 696 T; 0 U; 0 Other;

milarity 56.4%; Score 43.4; DB 1; Length 2928;

Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACATCACAGTCTCTGGGAGCATCATCTCTGCATGGTCTAGGTCATACTCCAAAC 60  
|||||  
GACAATCACAGTCTCTGGGAGCTGCCAAGCCTCCATCTCCAGCAACAATCCAAAC 1522  
|||||

CGTGGAGGACAGGAT 77

|||||

CGTGGAGGACAGGAT 1539

standard; cDNA; 2928 BP.

3 (revised)  
7 (first entry)

inoembryonic antigen cDNA.

cinoembryonic antigen; CEA; antibody; colon; cancer; ds.

ns.

Location/Qualifiers

1. .2097

/\*tag= a

1. .90

/\*tag= b

91. .2094

/\*tag= c

-A.

7.

7; 96JP-00221770.

7; 87JP-00006851.

NTORY LTD.

328482/30.

W22844.

inoembryonic antigen protein - used to prepare anti-CEA

page 9-12; 14pp; Japanese.

it sequence encodes human carcinoembryonic antigen (CEA), which d in the preparation anti-CEA antibodies (Ab). PolyA+ RNA was from the colon tissue of a colon cancer patient, and a lambda library prepared. Positive clones were separated by Ab

CC screening and analysed using various anti-CEA Ab. A cDNA libra CC prepared by the Okayama-Bark method. The cDNA clone, lambda Ki CC reacted with 4 anti-CEA Ab, was judged to encode human CEA. (U CC 25-MAR-2003 to correct PF field.)

XX Sequence 2928 BP; 825 A; 820 C; 588 G; 695 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 2; Length 2928;

Best Local Similarity 72.7%; Pred. No. 1.7e-05;

Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGCAATCACAGTCTCTGGGAGCATCATCTCTGCATGGTCTAGGTCATAAC

Db 1463 AGCAATCACAGTCTCTGGGAGCTGCCAAGCCTCCATCTCCAGCAAC

QY 61 CCCTGGAGGACAGGAT 77

Db 1523 CCCTGGAGGACAGGAT 1539

RESULT 39

ABL64746

ID ABL64746 standard; DNA; 2974 BP.

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

XX ABL64746;

cancer; ds; hepatocellular carcinoma; hepatotropic; liver tumour; cytostatic; expression profile; disease state; progression; drug toxicity; drug efficacy; drug metabolism.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

cleic search, using sw model

April 13, 2004, 17:11:06 ; Search time 1107.5 Seconds  
(without alignments)  
3013.466 Million cell updates/sec

US-10-090-326-24

1 agacaatcacagtctctgcg.....aacccgtggaggacaaggat 77

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

3470272 seqs, 21671516995 residues

hits satisfying chosen parameters: 6940544

length: 0

length: 2000000000

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hcg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_ov.\*

22: em\_or.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_btg\_hum.\*

31: em\_btg\_inv.\*

32: em\_btg\_other.\*

33: em\_btg\_mus.\*

34: em\_btg\_pln.\*

35: em\_btg\_rod.\*

36: em\_btg\_mam.\*

37: em\_btg\_vrt.\*

38: em\_sy.\*

39: em\_btgo\_hum.\*

40: em\_btgo\_mus.\*

41: em\_btgo\_other.\*

is the number of results predicted by chance to have a

score greater than or equal to the score of the result being  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Descrip
1	45	58.4	862	6	I08155	I08155
2	45	58.4	862	9	HUMCEALV	M17191
3	45	58.4	1623	6	AX497809	AX49780
4	45	58.4	2839	6	I08156	I08156
5	45	58.4	2839	6	I08166	I08166
6	43.4	56.4	111	6	AX260617	AX26061
7	43.4	56.4	190	6	AX260603	AX26060
8	43.4	56.4	326	6	AX261236	AX26123
9	43.4	56.4	326	6	AX261461	AX26146
10	43.4	56.4	326	6	AX261739	AX26173
11	43.4	56.4	571	6	AX261074	AX26107
12	43.4	56.4	572	6	AX261473	AX26147
13	43.4	56.4	1422	9	HUMCEAF	M16234
14	43.4	56.4	2019	6	AX505112	AX50511
15	43.4	56.4	2031	6	A39900	A39900
16	43.4	56.4	2031	6	I67748	I67748
17	43.4	56.4	2092	9	HSCRAASP	X16455
18	43.4	56.4	2097	6	A43169	A43169
19	43.4	56.4	2097	6	AR079553	AR07955
20	43.4	56.4	2106	6	AX133657	AX13365
21	43.4	56.4	2106	6	AX192349	AX19234
22	43.4	56.4	2106	6	AX393888	AX39388
23	43.4	56.4	2109	6	AX133977	AX13397
24	43.4	56.4	2109	6	AX468838	AX46883
25	43.4	56.4	2220	6	AR044683	AR04468
26	43.4	56.4	2349	6	AR052808	AR05280
27	43.4	56.4	2349	6	AR288121	AR28812
28	43.4	56.4	2434	6	AR052807	AR05280
29	43.4	56.4	2434	6	AR288120	AR28812
30	43.4	56.4	2888	9	BC034671	BC03467
31	43.4	56.4	2928	6	E01630	E01630
32	43.4	56.4	2928	6	E13123	E13123
33	43.4	56.4	2929	9	HUMCEA	M15042
34	43.4	56.4	2974	6	AX332574	AX33257
35	43.4	56.4	2974	6	AX409670	AX40967
36	43.4	56.4	2974	6	AX658324	AX65832
37	43.4	56.4	2974	6	AX677147	AX67714
38	43.4	56.4	2974	6	AX805532	AX80553
39	43.4	56.4	2974	9	HUMCEAF	M29540
40	43.4	56.4	3036	9	HUMANTCE	M17303
41	41	53.2	588	6	BD229772	BD22977
42	37.4	48.6	859	6	I08165	I08165
43	35	45.5	162661	2	AC146882	AC14688
44	35	45.5	193704	2	AC146845	AC14684
C 45	32.8	42.6	472	6	AR273307	AR27330

# ALIGNMENTS

RESULT 1	I08155	Sequence 1 from Patent EP 0346702.	DNA	linear	PA
LOCUS	I08155				
DEFINITION	I08155				
ACCESSION	I08155				
VERSION	I08155.1	GI:589126			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 862)				
AUTHORS	Barnett,T.R., Eiting,J.J., Kamarck,M.E. and Kretschnmer,				
TITLE	Transfected cell lines which express isolated genes of				
JOURNAL	carcinoembryonic antigen family				
	Patent: EP 0346702-A2 1 20-DEC-1989;				

Location/Qualifiers  
1. .862  
/organism="unknown"  
/mol\_type="unassigned DNA"

58.4%; Score 45; DB 6; Length 862;  
Similarity 74.0%; Pred. No. 6.1e-06;

Conservative 0; Mismatches 20; Indels 0; Gaps 0;

AGCATCAGTCTCTGGGAGGATCATCTCTGCATGGTCAGGTCAATCACTCAAAAC 60  
|||||  
AGCATCAGTCTCTGGGAGGATCATCTCTGCATGGTCAGGTCAATCACTCAAAAC 591

>GTGGAGGACCAAGGAT 77

>GTGGAGGACCAAGGAT 608

DEALV 862 bp mRNA linear PRI 01-NOV-1994  
in carcinoembryonic antigen mRNA, partial cds.

191

191.1 GI:180224

carcinoembryonic antigen.

; sapiens (human)

; sapiens

Chordata; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

(bases 1 to 862)

Brack, M.E., Elting, J.J., Hart, J.T., Goebel, S.J., Rae, P.M.,

Adurft, M.A., Nedwin, J.J., and Barnett, T.R.

carcinoembryonic antigen family: expression in a mouse L-cell

infectant and characterization of a partial cDNA in

peritoneal lambda gt11

2. Natl. Acad. Sci. U.S.A. 84 (15), 5350-5354 (1987)

30984

3415

final source text: Human carcinoma cell line LoVo (CCL229), cDNA  
clone lambda-clv7.

Location/Qualifiers

1. .862

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/map="19q13.2"

1. .862

/gene="CEA"

<1. >862

/notes="carcinoembryonic antigen"

/codon\_start=3

/protein\_id="AA051968.1"

/db\_xref="GI:180225"

/db\_xref="GDB:G00-119-054"

/translation="GLNHPKIPFTSNNSPVDEDAVALTCEPEIQNTYLVWVN

NQSLPVSRLQSGNDRITLLSVTRNDVGPYECGQNELSDVHSDPVLINLVIGDD

PTIPSYTYRPGVNLISLCHAASNPAPQYSLIDGNIQHTQELFISNITEKNSGLY

TCQANASGHSRTTKTIVSADVPKPSISSNNSPVDEDAVALTCEPEIQNTYLV

WVNGQSLPVSRLQSGNDRITLLSVTRNDVGPYECGQNELSDVHSDPVLINLVIGDD

GPDPPIISPPP"

58.4%; Score 45; DB 9; Length 862;  
Similarity 74.0%; Pred. No. 6.1e-06;

Conservative 0; Mismatches 20; Indels 0; Gaps 0;

GACATCAGTCTCTGGGAGGATCATCTCTGCATGGTCAGGTCAATCACTCAAAAC 60

|||||

GACATCAGTCTCTGGGAGGATCATCTCTGCATGGTCAGGTCAATCACTCAAAAC 591

|||||

>GTGGAGGACCAAGGAT 77

Db 592 CCGTGAGGACCAAGGAT 608  
|||||

RESULT 3

AX497809

LOCUS

AX497809

DEFINITION

Sequence 14 from Patent WO0240673.

ACCESSION

AX497809

VERSION

AX497809.1

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1

AUTHORS

Macina, R.A., Recipon, H., Chen, S.Y., Sun, Y., Liu, C. and

TITLE

Compositions and methods relating to lung specific genes

JOURNAL

Patent: WO 0240673-A 14 23-MAY-2002;

DIAGNOSIS

Diadex, Inc. (US)

FEATURES

Location/Qualifiers

1. .1623

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match

Best Local Similarity

58.4%; Score 45; DB 6; Length 1623;

Matches

57; Conservative 0; Mismatches 20; Indels 0;

QY 1 AGACATCAGTCTCTGGGAGGATCATCTCTGCATGGTCAGGTCAATCAAC

|||||

51 AGACATCAGTCTCTGGGAGGATCATCTCTGCATGGTCAGGTCAATCAAC

|||||

QY 61 CCGTGAGGACCAAGGAT 77

|||||

Db 111 CCGTGAGGACCAAGGAT 127

RESULT 4

I08156

LOCUS

I08156

DEFINITION

Sequence 2 from Patent EP 0346702.

ACCESSION

I08156

VERSION

I08156.1

KEYWORDS

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 2839)

AUTHORS

Barnett, T.R., Elting, J.J., Kamarch, M.E. and Kretschmer,

TITLE

Transfected cell lines which express isolated genes of

carcinoembryonic antigen family

JOURNAL

Patent: EP 0346702-A2 2 20-DEC-1989;

FEATURES

Location/Qualifiers

1. .2839

/organism="unknown"

/mol\_type="unassigned DNA"

ORIGIN

Query Match

Best Local Similarity

58.4%; Score 45; DB 6; Length 2839;

Matches

57; Conservative 0; Mismatches 20; Indels 0;

QY 1 AGACATCAGTCTCTGGGAGGATCATCTCTGCATGGTCAGGTCAATCAAC

|||||

Db 1479 AGACATCAGTCTCTGGGAGGATCATCTCTGCATGGTCAGGTCAATCAAC

|||||

QY 61 CCGTGAGGACCAAGGAT 77

|||||

Db 1539 CCGTGAGGACCAAGGAT 1555

```

166       2839 bp      DNA      linear      PAT 02-DEC-1994
quence 8 from Patent EP 0346710.
166
166.1 GI:589121
nown.
nown.
lassified.
(bases 1 to 2839)
net,T.R.D., Elting,J.J.D., Kamarck,M.E. and Kretscher,A.D.
As coding for members of the carcinoembryonic antigen family
ent: EP 0346710-A2 8 20-DEC-1989;
Location/Qualifiers
1. .2839
/organism="unknown"
/mol_type="unassigned DNA"

56.4%; Score 45; DB 6; Length 2839;
milarity 74.0%; Pred. No. 6.8e-06;
Conservative 0; Mismatches 20; Indels 0; Gaps 0;

GACAATCACAGTCTCTGGGAGCATCATCCTCTGCATGTCAGTCAACTCCAAAC 60
|||||
GACAATCACAGTCTCTGGGAGCATGTCCTCAAGCCCTCCATCTCCAGCAACTCCAAAC 1538
|||||
CGTGGAGGACAAGGAT 77
|||||
CGTGGAGGACAAGGAT 1555

50617       111 bp      DNA      linear      PAT 26-OCT-2001
ence 268 from Patent WO0173027.
50617
50617.1 GI:16509584
sapiens (human)
sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
alia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
her,M.J., Xu,J. and King,G.E.
ositions and methods for therapy and diagnosis of colon cancer
nt: WO 0173027-A 268 04-OCT-2001;
TXA CORPORATION (US)
Location/Qualifiers
1. .111
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

56.4%; Score 43.4; DB 6; Length 111;
milarity 72.7%; Pred. No. 2.1e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

TACAATCACAGTCTCTGGGAGCATCATCCTCTGCATGTCAGTCAACTCCAAAC 60
|||||
TACAATCACAGTCTCTGGGAGCTGCCCAAGCCCTCCATCTCCAGCAAACTCCAAAC 80
|||||
TGTGGAGGACAAGGAT 77
|||||
TGTGGAGGACAAGGAT 97

LOCUS      AX260603      190 bp      DNA      linear      PAT
DEFINITION Sequence 254 from Patent WO0173027.
ACCESSION  AX260603
VERSION     AX260603.1 GI:16509570
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Meagher,M.J., Xu,J. and King,G.E.
TITLE       Compositions and methods for therapy and diagnosis of c
JOURNAL     Patent: WO 0173027-A 254 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES    Location/Qualifiers
            1. .190
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match      56.4%; Score 43.4; DB 6; Length 190;
Best Local Similarity 72.7%; Pred. No. 2.2e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACAATCACAGTCTCTGGGAGCATCATCCTCTGCATGTCAGTCAACTCCAAAC
Db 21 AGACAATCACAGTCTCTGGGAGCTGCCCAAGCCCTCCATCTCCAGCAAAAC
QY 61 CGTGGAGGACAAGGAT 77
Db 81 CGTGGAGGACAAGGAT 97

RESULT 8
AX261236
LOCUS      AX261236      326 bp      DNA      linear      PAT
DEFINITION Sequence 887 from Patent WO0173027.
ACCESSION  AX261236
VERSION     AX261236.1 GI:16510203
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Meagher,M.J., Xu,J. and King,G.E.
TITLE       Compositions and methods for therapy and diagnosis of c
JOURNAL     Patent: WO 0173027-A 887 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES    Location/Qualifiers
            1. .326
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match      56.4%; Score 43.4; DB 6; Length 326;
Best Local Similarity 72.7%; Pred. No. 2.3e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACAATCACAGTCTCTGGGAGCATCATCCTCTGCATGTCAGTCAACTCCAAAC
Db 21 AGACAATCACAGTCTCTGGGAGCTGCCCAAGCCCTCCATCTCCAGCAAAAC
QY 61 CGTGGAGGACAAGGAT 77
Db 81 CGTGGAGGACAAGGAT 97

RESULT 9
AX261461
LOCUS      AX261461      326 bp      DNA      linear      PAT

```

ence 1112 from Patent WO0173027.  
1461  
1461.1 GI:16510428  
sapiens (human)  
yotia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
alia; Eutheria; Primates; Catarhini; Hominidae; Homo.  
her, M.J., Xu, J. and King, G.E.  
ositions and methods for therapy and diagnosis of colon cancer  
nt: WO 0173027-A 1112 04-OCT-2001;  
XA CORPORATION (US)  
Location/Qualifiers  
1..326  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
56.4%; Score 43.4; DB 6; Length 326;  
ilarity 72.7%; Pred. No. 2.3e-05;  
Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
ACATCATCAGTCTCTGGGAGCATCATCTCTGCATGGTCAGGTCTAATCCAAAC 60  
|||||  
ACATCATCAGTCTCTGGGAGCTGCCAAGCCCTCCATCTCCAGCAACACTCCAAAC 80  
|||||  
GTGGAGGACCAAGGAT 77  
|||||  
GTGGAGGACCAAGGAT 97  
|||||  
51739 326 bp DNA linear PAT 26-OCT-2001  
ence 1390 from Patent WO0173027.  
51739  
51739.1 GI:16510706  
sapiens (human)  
sapiens  
yotia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
alia; Eutheria; Primates; Catarhini; Hominidae; Homo.  
her, M.J., Xu, J. and King, G.E.  
ositions and methods for therapy and diagnosis of colon cancer  
nt: WO 0173027-A 1390 04-OCT-2001;  
XA CORPORATION (US)  
Location/Qualifiers  
1..326  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
56.4%; Score 43.4; DB 6; Length 326;  
ilarity 72.7%; Pred. No. 2.3e-05;  
Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
GACATCATCAGTCTCTGGGAGCATCATCTCTGCATGGTCAGGTCTAATCCAAAC 60  
|||||  
GACATCATCAGTCTCTGGGAGCTGCCAAGCCCTCCATCTCCAGCAACACTCCAAAC 80  
|||||  
CTGGAGGACCAAGGAT 77  
|||||  
CTGGAGGACCAAGGAT 97  
|||||  
61074 571 bp DNA linear PAT 26-OCT-2001  
ence 725 from Patent WO0173027.

ACCESSION AX261074 GI:16510041  
VERSION AX261074.1  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 Meagher, M.J., Xu, J. and King, G.E.  
AUTHORS Meagher, M.J., Xu, J. and King, G.E.  
TITLE Compositions and methods for therapy and diagnosis of c  
JOURNAL Patent: WO 0173027-A 725 04-OCT-2001;  
CORIXA CORPORATION (US)  
FEATURES Location/Qualifiers  
source 1..571  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 56.4%; Score 43.4; DB 6; Length 571;  
Best Local Similarity 72.7%; Pred. No. 2.4e-05;  
Matches 56; Conservative 0; Mismatches 21; Indels 0;  
QY 1 AGACAATCACAGTCTCTGGGAGCATCATCTCTGCATGGTCAGGTCTAATAAC  
|||||  
Db 21 AGACAATCACAGTCTCTGGGAGCTGCCAAGCCCTCCATCTCCAGCAACAAC  
|||||  
QY 61 CCGTGGAGGACCAAGGAT 77  
|||||  
Db 81 CCGTGGAGGACCAAGGAT 97  
|||||  
RESULT 12  
AX261473 572 bp DNA linear PAT  
LOCUS Sequence 1124 from Patent WO0173027.  
DEFINITION AX261473  
ACCESSION AX261473  
VERSION AX261473.1 GI:16510440  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 Meagher, M.J., Xu, J. and King, G.E.  
AUTHORS Meagher, M.J., Xu, J. and King, G.E.  
TITLE Compositions and methods for therapy and diagnosis of c  
JOURNAL Patent: WO 0173027-A 1124 04-OCT-2001;  
CORIXA CORPORATION (US)  
FEATURES Location/Qualifiers  
source 1..572  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 56.4%; Score 43.4; DB 6; Length 572;  
Best Local Similarity 72.7%; Pred. No. 2.4e-05;  
Matches 56; Conservative 0; Mismatches 21; Indels 0;  
QY 1 AGACAATCACAGTCTCTGGGAGCATCATCTCTGCATGGTCAGGTCTAATAAC  
|||||  
Db 21 AGACAATCACAGTCTCTGGGAGCTGCCAAGCCCTCCATCTCCAGCAACAAC  
|||||  
QY 61 CCGTGGAGGACCAAGGAT 77  
|||||  
Db 81 CCGTGGAGGACCAAGGAT 97  
|||||  
RESULT 13  
HUMCEAX 1422 bp mRNA linear PRI  
LOCUS Human carcinoembryonic antigen mRNA, 3' end.  
DEFINITION M16234  
ACCESSION



34.1 GI:180240  
 repeat; carcinoembryonic antigen.  
 sapiens (human)  
 ryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 alia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 bases 1 to 1422)  
 ermann,W., Ortlieb,B., Friedrich,R. and von Kleist,S.  
 action and characterization of cDNA clones encoding the human  
 inoembryonic antigen reveal a highly conserved repeating  
 cture  
 . Natl. Acad. Sci. U.S.A. 84 (9), 2960-2964 (1987)  
 4247  
 671  
 inal source text: Human colon adenocarcinoma, cDNA to mRNA,  
 es pCEA[1-5].  
 t entry and printed copy of sequence [1] kindly provided by  
 mmermann, 01-JUL-1987.  
 Location/Qualifiers  
 1..1422  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /map="19q13.2"  
 1..1422  
 /gene="CEA"  
 /gene="CEA"  
 <1..1422  
 /product="carcinoembryonic antigen mRNA"  
 <1..1121  
 /gene="CEA"  
 /note="carcinoembryonic antigen"  
 /codon\_start=2  
 /protein\_id="AAA51972.1"  
 /db\_xref="GI:180241"  
 /db\_xref="GDB:G00-119-054"  
 /translation="NSNPVEDEDAVALTCEPEIQNTTTLVWVNNQSLVPSRLQSLND  
 NRITLTSVTRNDVGPVECGQLNLSVDHSDPVLINLVYGDPTISSTYTRPGVN  
 LSLSCHAASNPQAQYSLIDNIGNQHTQELFISNTEKNSGLYTCQANNSAGHSRIT  
 VKTITVAELPKPSISNNKSPVEDKDAVFTCEPEACNTTLYLWVNGSLPVSRLQ  
 LSGNRITLTFNVTRNDARAYVCGIQSVAANRSDPVTLDVLYGPDTPITISPPSSYL  
 SGANLNLSCHASNPSPQYSWRINGIPQHTQVLFIAKITNNNGTVACFVSNLATGR  
 NNSIVKITSVTSAGTSPGLSAGATVGMIGVLGVALLI"  
 on  
 <1..502  
 /note="repeat copy A"  
 on 503..1036  
 /note="repeat copy B"  
 on 1241..>1422  
 /note="Alu repeat"  
 bp upstream of PstI site.  
 56.4%; Score 43.4; DB 9; Length 1422;  
 ilarity 72.7%; Pred. No. 2.6e-05;  
 Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
 ACAATCACAGTCTCTCGGAGCATCATCTCTGCATGTCAGTCATCACTCAAC 60  
 |||||  
 ACAATCACAGTCTCTCGGAGTGCACAGCCCTCATCTCCAGCAACCACTCAAC 545  
 |||||  
 GTGGAGGACAGGAT 77  
 |||||  
 GTGGAGGACAGGAT 562  
 |||||  
 5112 2019 bp DNA linear PAT 27-SEP-2002  
 ence 5 from Patent WO0240059.  
 5112  
 5112.1 GI:23386419  
 sapiens (human)  
 sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hom  
 1  
 Mincheff,M.S., Loukinov,D.I. and Zoubak,S.  
 Methods and compositions for inducing cell-mediated immu  
 responses  
 Patent: WO 0240059-A 5 23-MAY-2002;  
 Alenco S. (US) ; Loukinov, Dmitri I. (US) ; Zoubak, Serg  
 Location/Qualifiers  
 1..2019  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 7..2019  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD48803.1"  
 /db\_xref="GI:23386420"  
 /db\_xref="REMBL:CAD48803"  
 /translation="MAKLTIESTPFNVAGKEVLLVHNLPHLFGY  
 RQIIGVIGTQATPGPAYSGREIIVPNASLLIQNDTGFVTLH  
 TQGFVYVPELPKPSISNNKSPVEDKDAVFTCEPEQDQATYLVWVN  
 LSGNRITLTFNVTRNDTASYKCTONPVARSDSVLNVLGPD  
 SGENLNLSCHAASNPQAQYSWFVNGTQQSTQELFIPNITVNSGSY  
 NRITVITVYAEPPKPFITSNNSNPVEDEDAVLTCEPEIQNTTLY  
 PRIQLSNDNRITLTSVTRNDVGPVECGIQNELSVDSHSDPVLINLVY  
 TYYPGVNLSLSCHAASNPQAQYSLIDNIGNQHTQELFISNTEKN  
 ASGSRITVKTITVSAELPKPSISNNKSPVEDKDAVFTCEPEAQN  
 LPVSPRLQSLNNGNRITLTFNVTRNDARAYVCGIQSVAANRSDPVT  
 SPDSVSYLSGANLNLSCHASNPSPQYSWRINGIPQHTQVLFIAKI  
 VSNLATGRNNSIVKITSVTSAGTSPGLSAGATVGMIGVLGVALLI"  
 ORIGIN  
 Query Match 56.4%; Score 43.4; DB 6; Length 2019;  
 Best Local Similarity 72.7%; Pred. No. 2.7e-05;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0;  
 Qy 1 AGCAATCACAGTCTCTCGGAGCATCATCTCTGCATGTCAGTCATAACT  
 |||||  
 Db 1385 AGCAATCACAGTCTCTCGGAGTGCACAGCCCTCATCTCCAGCAACCACT  
 |||||  
 Qy 61 CCGTGGAGGACAGGAT 77  
 |||||  
 Db 1445 CCGTGGAGGACAGGAT 1461  
 |||||  
 RESULT 15  
 A39900  
 LOCUS A39900 2031 bp DNA linear PAT  
 DEFINITION Sequence 2 from Patent EP0618292.  
 ACCESSION A39900  
 VERSION A39900.1 GI:2296118  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hom  
 1 (bases 1 to 2031)  
 Terakikh,A., Pelegrin,A. and Mach,J.  
 CEA derivatives lacking the Hydrophobic C-terminal domai  
 Patent: EP 0618292-A 2 05-OCT-1994;  
 HOFFMANN LA ROCHE (CH)  
 Other publication CA 2116640 940926  
 Other publication JP 6321992 941122.  
 FEATURES  
 source  
 1..2031  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 56.4%; Score 43.4; DB 6; Length 2031;

ilarity 72.7%; Pred. No. 2.7e-05; Indels 0; Gaps 0;  
Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
ACAATCACAGTCTCTGCGGAGCATCATCTCTGCGATGTCAGGTCACTAACTCAAAAC 60  
|||||  
ACAATCACAGTCTCTGCGGAGCTGCCAAGCCCTCCATCTCCAGCACTCAAAAC 1534  
GTGAGGACAAAGGAT 77  
GTGAGGACAAAGGAT 1551  
48 2031 bp DNA linear PAT 30-DEC-1997  
ence 2 from patent US 5672513.  
48 2031 bp DNA linear PAT 30-DEC-1997  
48.1 GI:2731283  
DOWN.  
assified.  
bases 1 to 2031)  
kikh,A., Pellegrin,A. and Mach,J.-P.  
ino-embryonic antigen derivatives lacking the carboxyl terminal  
nt: US 5672513-A 2 30-SEP-1997;  
Location/Qualifiers  
1. .2031  
/organism="unknown"  
/mol\_type="unassigned DNA"  
56.4%; Score 43.4; DB 6; Length 2031;  
ilarity 72.7%; Pred. No. 2.7e-05;  
Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
ACAATCACAGTCTCTGCGGAGCATCATCTCTGCGATGTCAGGTCACTAACTCAAAAC 60  
|||||  
ACAATCACAGTCTCTGCGGAGCTGCCAAGCCCTCCATCTCCAGCACTCAAAAC 1534  
XGTGAGGACAAAGGAT 77  
XGTGAGGACAAAGGAT 1551  
3AASP 2092 bp mRNA linear PRI 19-JUL-1995  
in mRNA for carcinoembryonic antigen pCEA80-11.  
155.1 GI:29854  
arnative splicing; carcinoembryonic antigen; cea gene; cell  
sion molecule; cell surface protein.  
> sapiens (human)  
> sapiens  
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
nalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
(bases 1 to 2092)  
awa.S.  
act Submission  
mitted (02-SEP-1995) Oikawa S., Suntory Institute for Biomedical  
earch, 1-1-1 Wakayamadai, Shimamoto-cho, Mishima-gun Osaka 618,  
an  
hor describes alternative splicing which leads to a three base  
r deletion (gca) in this CEA mRNA.  
also <M15042> for sequence of CEA mRNA.  
Location/Qualifiers  
1. .2092  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="colonic mucosa"

CDS  
1. .>2092  
/codon\_start=2  
/product="pCEA80-11 protein (647 AA)"  
/protein\_id="CAA34474.1"  
/db\_xref="GI:825638"  
/db\_xref="GOA:P06731"  
/db\_xref="SWISS-PROT:P06731"  
/translation="SAPHRWCIPQWRLLLTASLLTFWNPPTAKLI  
KEVLLVHNLFPQHLFGYSWYKGRVGNRQIIGYVIGTQQATPGPAI  
LLIQNIQNDTGYTLHVIKSDLVNEEATGQFVYVPELPKPSISSNI  
FTCEPTQDATYLLWVNNQSLVSPRLQSLNGNRTLTLFNVTRNDT  
ARRSDSVILANVLYGPDAPITPLNTSYRSGENLNLSCHAAASNPAPQI  
TOELFNIITVNSGSGYTCQAHNSDTGLNRTVTITIVYEPKPFII  
AVALTCEPEIQNTYLLWVNNQSLVSPRLQSLNDNRTLTLVSTRN  
ELSDVHSDPVLNLYGDDPTISPSYTYRPGVNLSSLCHAAASNP  
QQHTELEFIINITEKNKSLYTCQANNSAGHSRTVTKITITVSAELPF  
EDKDAVATCEPEAQNTYLLWVNNQSLVSPRLQSLNGNRTLTLF  
GIQNSVANSRSDPVTLDVLYGPDPTIISPPDSSYLSGANLNLSCHS  
NGIPQOHTOVLFIAKITPNNNGTYACFVSNLATGRNNSIVKSIIVS  
TVGIMIGVLGVALL"  
947. .948  
/note="site of 3 bp (gca) deletion"  
misc\_feature  
ORIGIN  
Query Match 56.4%; Score 43.4; DB 9; Length 2092;  
Best Local Similarity 72.7%; Pred. No. 2.7e-05;  
Matches 56; Conservative 0; Mismatches 21; Indels 0;  
QY 1 AGACAATCACAGTCTCTGCGGAGCATCATCTCTGCGATGTCAGGTCACTAACT  
DB 1461 AGACRATCACAGTCTCTGCGGAGCTGCCAAGCCCTCCATCTCCAGCACT  
QY 61 CCGTGGAGGACAAAGGAT 77  
DB 1521 CCGTGGAGGACAAAGGAT 1537  
RESULT 18  
A43169 2097 bp DNA linear PAT  
LOCUS Sequence 35 from Patent WO9506067.  
DEFINITION A43169  
ACCESSION A43169  
VERSION A43169.1 GI:2298545  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hor  
Bodmer,W.F., Durbin,H., Smayda,D., Stewart,L.M., Young,S  
Bates,P.A.  
MONOCLONAL ANTIBODIES FOR USE IN DIAGNOSIS AND TREATMEN  
COLORECTAL CANCER  
Patent: WO 9506067-A 35 02-MAR-1995;  
IMP CANCER RES TECH (GB)  
Other publication CA 2168440 950302  
Other publication AU 7390694 950321.  
FEATURES  
source  
1. .2097  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
CDS  
1. .2094  
/note="unnamed protein product; Protein sequen  
conflict with the conceptual translation"  
/codon\_start=1  
/protein\_id="CAA02706.1"  
/db\_xref="GI:2298546"  
/translation="SAPHRWCIPQWRLLLTASLLTFWNPPTAKLI  
KEVLLVHNLFPQHLFGYSWYKGRVGNRQIIGYVIGTQQATPGPA  
LLIQNIQNDTGYTLHVIKSDLVNEEATGQFVYVPELPKPSISSNI  
FTCEPTQDATYLLWVNNQSLVSPRLQSLNGNRTLTLFNVTRNDT

ARRSDSVILNVLYGPDAPTISPLNTSYRSGENLNLSCAASNPPAQYWFVNGTQQS  
TOELFIBNTVNNSSYTCQAHNSDTGLNRTVTITVYAEPPKPFITSNNSNPVEDE  
DAVATCEPFIQNTIYLMWVNNQSLPVSRLQSLNDNRLLTLLSVTRNDVGYECGIQ  
NELSDVHSDPVLNVLYGDDPTISPSYTYRPGVNLNLSCHAASNPPAQYSLNLDGN  
IQOQTOELFISNITEKNSGLYTCQAHNSGHSRTTAKTITVSAELPKPSISSNLSKP  
VEDDAVAFCEPQAQNTIYLMWVNNQSLPVSRLQSLNDNRLLTLLSVTRNDARAYV  
CGIQNSVANRSDPVLNVLYGDDPTIISPPDSSYLSGANLNLSCAASNPPAQYSWR  
INGIPOQHTOVLPIAKITPNNGTYACFVSNLATGRNNSIVKSIIVSASGSPGLSAG  
ATVGMIGVLGVALLI"

56.4%; Score 43.4; DB 6; Length 2097;  
ilarity 72.7%; Pred. No. 2.7e-05;  
Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
ACAAATCACAGTCTCTGCGGAGCAAGGAT 77  
|||||  
ACAAATCACAGTCTCTGCGGAGCAAGGAT 1539  
|||||  
GTGGAGGACAAAGGAT 77  
|||||  
GTGGAGGACAAAGGAT 1539

9553  
ence 35 from patent US 5965710.  
9553  
9553.1 GI:10006297  
own.  
own.  
assified.  
bases 1 to 2097)  
er,W.F., Durbin,H., Snary,D., Stewart,L.M.D., Young,S. and  
s,P.A.  
clonal antibodies for use in diagnosis and treatment of  
rectal cancer  
nt: US 5965710-A 35 12-OCT-1999;  
Location/Qualifiers  
1. .2097  
/organism="unknown"  
/mol\_type="unassigned DNA"

56.4%; Score 43.4; DB 6; Length 2097;  
ilarity 72.7%; Pred. No. 2.7e-05;  
Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
ACAAATCACAGTCTCTGCGGAGCAAGGAT 77  
|||||  
ACAAATCACAGTCTCTGCGGAGCAAGGAT 1539  
|||||  
GTGGAGGACAAAGGAT 77  
|||||  
GTGGAGGACAAAGGAT 1539

3657  
ence 111 from Patent WO0130382.  
3657  
3657.1 GI:14139699  
hetic construct  
hetic construct  
ficial sequences.  
natein,N., Tartaglia,J., Moingeon,P. and Barber,B.  
od of inducing and/or enhancing an immune response to tumor  
gens

JOURNAL Patent: WO 0130382-A 111 03-MAY-2001;  
Aventis Pasteur Limited (CA)  
Location/Qualifiers  
source  
1. .2106  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
CDS  
1. .2106  
/note="unnamed protein product; modified CEA"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC38955.1"  
/db\_xref="GI:14139700"  
/db\_xref="REMTREMBL:CAC38955"  
/translation="MESPSAPPHRWCIPWQRLLLTASLLTFWNPPT  
VAEGKEVLLVHNLPHQLFQHSYWKYGERVDGNRQIIGYVIGTQAT  
PNASILLIQNDTGFTYTHVKSIDLVNEBATQFRVYBELPKPS  
DAVAFCEPQTDATYLMWVNNQSLPVSRLQSLNDNRLLTLLSVTR  
NPVSARRSDSVILNVLYGPDAPTISPLNTSYRSGENLNLSCAASN  
FQOQTOELFISNITEKNSGLYTCQAHNSDTGLNRTVTITVYBPP  
EDEDAVAFCEPQNTIYLMWVNNQSLPVSRLQSLNDNRLLTLLSV  
GIONELSDVHSDPVLNVLYGDDPTISPSYTYRPGVNLNLSCHA  
DQNTIQOHTOELFISNITEKNSGLYTCQAHNSGHSRTTAKTITV  
SKPVEDDAVAFCEPQAQNTIYLMWVNNQSLPVSRLQSLNDNR  
AYVCGIQNSVANRSDPVLNVLYGDDPTIISPPDSSYLSGANLNL  
SWRINGIPQOHTOVLPIAKITPNNGTYACFVSNLATGRNNSIVKS  
SAGATVGMIGVLGVALLI"

ORIGIN  
Query Match 56.4%; Score 43.4; DB 6; Length 2106;  
Best Local Similarity 72.7%; Pred. No. 2.7e-05;  
Matches 56; Conservative 0; Mismatches 21; Indels 0;  
QY 1 AGACAATCACAGTCTCTGCGGAGCAAGGAT 77  
Db 1472 AGACAATCACAGTCTCTGCGGAGCAAGGAT 1548  
QY 61 CCGTGGAGGACAAAGGAT 77  
Db 1532 CCGTGGAGGACAAAGGAT 1548

RESULT 21  
AXI92349  
LOCUS AXI92349 2106 bp DNA linear PAT  
DEFINITION Sequence 3 from Patent WO0149317.  
ACCESSION AXI92349  
VERSION AXI92349.1 GI:15210326  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Emage, P., Barber, B.H., Sambhara, S. and Sia, C.D.  
TITLE Enhancing the immune response to an antigen by presensit  
an inducing agent prior to immunizing with the inducing  
the antigen  
JOURNAL Patent: WO 0149317-A 3 12-JUL-2001;  
Aventis Pasteur Limited (CA)  
Location/Qualifiers  
source  
1. .2106  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
CDS  
1. .2106  
/note="unnamed protein product; modified CEA"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC51319.1"  
/db\_xref="GI:15210327"  
/db\_xref="REMTREMBL:CAC51319"  
/translation="MESPSAPPHRWCIPWQRLLLTASLLTFWNPPT  
VAEGKEVLLVHNLPHQLFQHSYWKYGERVDGNRQIIGYVIGTQAT

PNASLIIQNIQNDTGFTYLVKSDLVNEBATGQFRVTPPELKPISNNKSPVEDK  
 DVAFTCEPETQDATYLVWNNQSLVSPRLQSLNGNRTLLTFNVRNTASYKCBTQ  
 NPVSARSDVILNVLYGDPTISPLNTSYRSGENILNSCHAAFNPPAQYFWNGT  
 FQOSTELFIPNITVNNSSYTCQAHNSDTGLNRTTITVVEPKPTITNSNPV  
 EDEDAVACEPIQNTYLVWNNQSLVSPRLQSLNDNRITLLSVTRNDVGPYEC  
 GICNELSDHSDPVLNVLYGDDPTISPSYTYRSGVNLISLCHAASNPAPQYSMLI  
 DGNIOHTQELFISNTERKNSGLYTCQAHNSASGHSRTTKLITVSAELPKFSISNN  
 SKPEVDKDAVFTCEPEAQNTYLVWNNQSLVSPRLQSLNGNRTLLTFNVRNTDAR  
 AYVCGIQNSVANSRSDPVLNVLYGDDPTISPSYTYRSGVNLISLCHAASNPAPQY  
 SWRINGIPQOHTQVLFIAKITPNNGTYACFVSNLATGRNNSIVKSIIVSASGTSPL  
 SAGATVGMIGVLGVGVALI"

56.4%; Score 43.4; DB 6; Length 2106;

ilarity 72.7%; Pred. No. 2.7e-05; Gaps 0;  
 Conservative 0; Mismatches 21; Indels 0;

3ACAATCACAGTCTCTGCGAGAGCATCATCTCTGCGAGTTCAGGTTCATTAACCTCAAAAC 60  
 |||||  
 3ACAATCACAGTCTCTGCGAGAGTCCCAAGCCCTCCATCTCCAGCAACACTCCAAAC 1531

CGTGAGGACAAGGAT 77  
 |||||  
 CGTGAGGACAAGGAT 1548

93888 2106 bp DNA linear PAT 23-MAR-2002  
 nence 2 from Patent WO0210379.

93888  
 93888.1 GI:19701852

thetic construct  
 thetic construct  
 ificial sequences.

instein,N., Tartaglia,J., Tine,J.A., Panicali,D.L., Gritz,J. and  
 lom,J.  
 ified cea and uses thereof  
 ent: WO 0210379-A 2 07-FEB-2002;  
 ntis Pasteur Limited (CA); Therion Biologics (US); National  
 cer Institute (US)

Location/Qualifiers  
 1..2106  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Cea modified polypeptide"

56.4%; Score 43.4; DB 6; Length 2106;

ilarity 72.7%; Pred. No. 2.7e-05; Gaps 0;  
 Conservative 0; Mismatches 21; Indels 0;

3ACAATCACAGTCTCTGCGAGAGCATCATCTCTGCGAGTTCAGGTTCATTAACCTCAAAAC 60  
 |||||  
 3ACAATCACAGTCTCTGCGAGAGTCCCAAGCCCTCCATCTCCAGCAACACTCCAAAC 1531

CGTGAGGACAAGGAT 77  
 |||||  
 CGTGAGGACAAGGAT 1548

33977 2109 bp DNA linear PAT 15-MAY-2001  
 nence 4 from Patent WO0124832.

33977  
 33977.1 GI:14139917

thetic construct  
 thetic construct

artificial sequences.  
 1  
 Pecher,G.  
 Pharmaceutical composition for treating and preventing  
 tumors, which express the tumor antigen mucin and/or th  
 carcinoembryonic antigen (ce a), and the use thereof  
 Patent: WO 0124832-A 4 12-APR-2001;

Pecher, Gabriele (DE)  
 Location/Qualifiers  
 1..2109

/organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Gen des humanen Carcinoembryonalen Anti  
 1..2109  
 /note="unnamed protein product"

/codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAC38973.1"

/db\_xref="GI:14139918"  
 /translation="MESPSAPPHRWICIPQWRLLLTASLITFWNPPT  
 VAEKVELLVHNLPHLFGYSWKYGERVDGNRQIIGYVIGTQOAT  
 PNASLIIQNIQNDTGFTYLVKSDLVNEBATGQFRVTPPELKPES  
 DVAFTCEPETQDATYLVWNNQSLVSPRLQSLNGNRTLLTFNVT  
 NPVSARSDVILNVLYGDPTISPLNTSYRSGENILNSCHAAASN  
 FQOSTELFIPNITVNNSSYTCQAHNSDTGLNRTTITVVEPKPT  
 VEDEDAVACEPIQNTYLVWNNQSLVSPRLQSLNDNRITLLSVTR  
 CGICNELSDHSDPVLNVLYGDDPTISPSYTYRSGVNLISLCHA  
 IDGNIOHTQELFISNTERKNSGLYTCQAHNSASGHSRTTKLITV  
 NSKPEVDKDAVFTCEPEAQNTYLVWNNQSLVSPRLQSLNGNR  
 RAYVCGIQNSVANSRSDPVLNVLYGDDPTISPSYTYRSGVNLIS  
 YSWRINGIPQOHTQVLFIAKITPNNGTYACFVSNLATGRNNSIVK  
 LSAGATVGMIGVLGVGVALI"

## ORIGIN

Query Match 56.4%; Score 43.4; DB 6; Length 2109;  
 Best Local Similarity 72.7%; Pred. No. 2.7e-05;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACAATCACAGTCTCTGCGAGAGCATCATCTCTGCGAGTTCAGGTTCATTAACCTCAAAAC  
 Db 1475 AGACAATCACAGTCTCTGCGAGAGTCCCAAGCCCTCCATCTCCAGCAACAAAC

QY 61 CGTGAGGACAAGGAT 77

Db 1535 CGTGAGGACAAGGAT 1551

## RESULT 24

AX468838 2109 bp DNA linear PAT  
 Sequence 12 from Patent WO0222803.

AX468838

VERSION AX468838.1 GI:21901578

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho

## REFERENCE

1  
 AUTHORS Chicz,R. and Tomlinson,A.  
 TITLE Peptide epitopes recognized by antigen specific cd4 + t  
 JOURNAL Patent: WO 0222803-A 12 21-MAR-2002;  
 FEATURES Zycos Inc. (US)

## source

Location/Qualifiers  
 1..2109  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

Query Match 56.4%; Score 43.4; DB 6; Length 2109;

ilarity 72.7%; Pred. No. 2.7e-05;  
 Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
 AGAATACAGTCTCTGGAGAGCATATCCTCTGATGGTCAGGTCAATCACTCAAAAC 60  
 |||||  
 AGAATACAGTCTCTGGAGAGCATATCCTCTGATGGTCAGGTCAATCACTCAAAAC 1534  
 |||||  
 CGTGAGGACAAAGGAT 77  
 |||||  
 CGTGAGGACAAAGGAT 1551  
 |||||  
 44683 2220 bp DNA linear PAT 29-SEP-1999  
 44683 16 from patent US 5817512.  
 44683.1 GI:5966148  
 LOCUS  
 DEFINITION 2220 bp DNA linear PAT 29-SEP-1999  
 ACCESSION AR288121  
 VERSION AR288121.1 GI:5966148  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 2220)  
 AUTHORS Paoletti, E., Tartaglia, J., and Cox, W. I.  
 TITLE Vaccinia virus expressing cytokine and/or tumor associated genes  
 JOURNAL Patent: US 5817512-A 16 OCT-1998;  
 FEATURES  
 source Location/Qualifiers  
 1..2220  
 /organism="unknown"  
 /mol\_type="unassigned DNA"  
 ORIGIN  
 Query Match 56.4%; Score 43.4; DB 6; Length 2349;  
 Best Local Similarity 72.7%; Pred. No. 2.7e-05;  
 Mismatches 0; Mismatches 21; Indels 0; Gaps 0;  
 AGAATACAGTCTCTGGAGAGCATATCCTCTGATGGTCAGGTCAATCACTCAAAAC 60  
 |||||  
 AGAATACAGTCTCTGGAGAGCATATCCTCTGATGGTCAGGTCAATCACTCAAAAC 1630  
 |||||  
 GTGAGGACAAAGGAT 77  
 |||||  
 GTGAGGACAAAGGAT 1647  
 |||||  
 12808 2349 bp DNA linear PAT 29-SEP-1999  
 12808 145 from patent US 5833975.  
 12808.1 GI:5977670  
 LOCUS  
 DEFINITION 2349 bp DNA linear PAT 29-SEP-1999  
 ACCESSION AR052807  
 VERSION AR052807.1 GI:5977670  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 2349)  
 AUTHORS Paoletti, E., Tartaglia, J., and Cox, W. I.  
 TITLE Vaccinia virus expressing cytokine and/or tumor-associated antigen sequence  
 JOURNAL Patent: US 5833975-A 14 NOV-1998;  
 FEATURES  
 source Location/Qualifiers  
 1..2349  
 /organism="unknown"  
 /mol\_type="unassigned DNA"  
 ORIGIN  
 Query Match 56.4%; Score 43.4; DB 6; Length 2349;  
 Best Local Similarity 72.7%; Pred. No. 2.7e-05;  
 Mismatches 0; Mismatches 21; Indels 0; Gaps 0;  
 AGAATACAGTCTCTGGAGAGCATATCCTCTGATGGTCAGGTCAATCACTCAAAAC 60  
 |||||  
 AGAATACAGTCTCTGGAGAGCATATCCTCTGATGGTCAGGTCAATCACTCAAAAC 1630  
 |||||

Db 1658 AGAATACAGTCTCTGGAGAGCATATCCTCTGATGGTCAGGTCAATCACTCAAAAC  
 QY 61 CGTGAGGACAAAGGAT 77  
 |||||  
 Db 1718 CGTGAGGACAAAGGAT 1734  
 |||||  
 RESULT 27  
 AR288121 2349 bp mRNA linear PAT  
 LOCUS AR288121  
 DEFINITION Sequence 145 from patent US 6537594.  
 ACCESSION AR288121  
 VERSION AR288121.1 GI:31675400  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 2349)  
 AUTHORS Paoletti, E., Tartaglia, J., and Cox, W. I.  
 TITLE Vaccinia virus comprising cytokine and/or tumor associated genes  
 JOURNAL Patent: US 6537594-A 145 25-MAR-2003;  
 FEATURES  
 source Location/Qualifiers  
 1..2349  
 /organism="unknown"  
 /mol\_type="mRNA"  
 ORIGIN  
 Query Match 56.4%; Score 43.4; DB 6; Length 2349;  
 Best Local Similarity 72.7%; Pred. No. 2.7e-05;  
 Mismatches 56; Conservative 0; Mismatches 21; Indels 0;  
 AGAATACAGTCTCTGGAGAGCATATCCTCTGATGGTCAGGTCAATCACTCAAAAC 60  
 |||||  
 AGAATACAGTCTCTGGAGAGCATATCCTCTGATGGTCAGGTCAATCACTCAAAAC 1734  
 |||||  
 CGTGAGGACAAAGGAT 77  
 |||||  
 CGTGAGGACAAAGGAT 1734  
 |||||  
 RESULT 28  
 AR052807 2434 bp DNA linear PAT  
 LOCUS AR052807  
 DEFINITION Sequence 144 from patent US 5833975.  
 ACCESSION AR052807  
 VERSION AR052807.1 GI:5977669  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 2434)  
 AUTHORS Paoletti, E., Tartaglia, J., and Cox, W. I.  
 TITLE Vaccinia virus expressing cytokine and/or tumor-associated antigen sequence  
 JOURNAL Patent: US 5833975-A 144 10-NOV-1998;  
 FEATURES  
 source Location/Qualifiers  
 1..2434  
 /organism="unknown"  
 /mol\_type="unassigned DNA"  
 ORIGIN  
 Query Match 56.4%; Score 43.4; DB 6; Length 2434;  
 Best Local Similarity 72.7%; Pred. No. 2.7e-05;  
 Mismatches 56; Conservative 0; Mismatches 21; Indels 0;  
 AGAATACAGTCTCTGGAGAGCATATCCTCTGATGGTCAGGTCAATCACTCAAAAC 60  
 |||||  
 AGAATACAGTCTCTGGAGAGCATATCCTCTGATGGTCAGGTCAATCACTCAAAAC 1731  
 |||||  
 CGTGAGGACAAAGGAT 77  
 |||||  
 CGTGAGGACAAAGGAT 1731  
 |||||

88120  
 pence 144 from patent US 6537594.  
 :88120 GI:31675399

known.  
 classified.  
 (bases 1 to 2434)  
 lerti, E., Tartaglia, J. and Cox, W.I.  
 cina virus comprising cytokine and/or tumor associated antigen  
 es

ent: US 6537594-A 144 25-MAR-2003;  
 Location/Qualifiers  
 1. .2434  
 /organism="unknown"  
 /mol\_type="mRNA"

56.4%; Score 43.4; DB 6; Length 2434;  
 milarity 72.7%; Pred. No. 2.7e-05;  
 Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCGCGAAGCATCATCTCTGCATGGTCAAGTCATCACTCAAAAC 60  
 GACAATCACAGTCTCGCGAAGCATCATCTCTGCATGGTCAAGTCATCACTCAAAAC 1714

CGTGAGGACAGGAT 77  
 |||||  
 CGTGAGGACAGGAT 1731

34671 2888 bp mRNA linear PRI 07-OCT-2003  
 o sapiens carcinoembryonic antigen-related cell adhesion  
 ecule 5, mRNA (CDNA clone MGC:34212 IMAGE:5184800), complete

34671 GI:21961633  
 o sapiens (human)  
 o sapiens  
 aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 malia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 (bases 1 to 2888)  
 ausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
 usner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 schul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
 kins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
 tchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 pleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 eetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
 ninci, P., Prange, C., Raba, S.S., Loquellano, N.A., Peters, G.J.,  
 anson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
 ernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 ley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
 lalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 ey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
 chez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 lffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 kson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 terfield, J.S., Krzywinski, M.I., Skalska, J., Smalios, D.E.,  
 merch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
 eration and initial analysis of more than 15,000 full-length  
 an and mouse cDNA sequences

U.S. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 88257  
 77932

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

2 (bases 1 to 2888)  
 Strausberg, R.  
 Direct Submission  
 Submitted (24-JUL-2002) National Institutes of Health,  
 Gene Collection (MGC), Cancer Genomics Office, National  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20  
 USA

REMARK  
 COMMENT

NIH-MGC Project URL: http://mgc.nci.nih.gov  
 Contact: MGC help desk  
 Email: gqaps-x@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILM)  
 DNA Sequencing by: Baylor College of Medicine Human Ge  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
 Contact: amg@bcm.tmc.edu  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lo  
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Ne  
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information  
 through the I.M.A.G.E. Consortium/LLNL at: http://image  
 Series: IRAK Plate: 50 Row: e Column: 17  
 This clone was selected for full length sequencing beca  
 passed the following selection criteria: matched mRNA s  
 Location/Qualifiers  
 1. .2888  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:34212 IMAGE:5184800"  
 /tissue\_type="Colon, Kidney, Stomach, adult, w  
 /clone\_lib="NIH MGC\_116"  
 /lab\_host="DH10E"  
 /note="Vector: pCMV-SPORT6"

1. .2888  
 /genes="CEACAM5"  
 /note="synonyms: CEACAM5, CD66e"  
 /db\_xref="LocusID:1048"  
 /db\_xref="MIM:114890"  
 96. .2204  
 /codon\_start=1  
 /product="CEACAM5 protein"  
 /protein\_id="AAH34671.1"  
 /db\_xref="GI:21961634"  
 /db\_xref="LocusID:1048"  
 /translation="MESPAPPHRCWICPQRLLLTASLLTFWNPPI  
 VAGKEVLLVHNPQHLFGYSWKGVGRVGNRQIIGYVIGTQOAI  
 ENASLLQIONDTGFTLVHTKSLVNEEATGQPRVPELPKPS  
 DAVFTCEBETQATYLVWVNNQSLPVSPLQNSGNRLITLNFVI  
 NPVSARRSDVILNVLYGDPAPTISPLNTSYRSGENLSCHAAAN  
 FOQSTQELFPNITVNNSSYTCQAHNSDTGLNRTVTITVYAEF  
 VEDDAVALTCPEIQNTYLVWVNNQSLPVSPLQNSGNRLITL  
 CTGNELSDVSDPVLNVLYGDPDPTISPSYTYRPGYNLSLSCH  
 IGNICQHTQELFISNITKNSGLYTCQANNSAGHSRTVTITV  
 NKPVEDKDAVFTCEBETQATYLVWVNNQSLPVSPLQNSGNR  
 RAYVQIQNSVSNRSDPVLVLYGDPDPTISPSYTYRPGYNLSCHAAAN  
 YSWRINQHTQELFISNITKNSGLYTCQANNSAGHSRTVTITV  
 LSGATVGMIGVILVGVALLI"

564. .755  
 /note="IG; Region: Immunoglobulin"  
 /db\_xref="CDD:smart00409"

825. .1046  
 /note="IG; Region: Immunoglobulin"  
 /db\_xref="CDD:smart00409"

1110. .1331  
 /note="IG; Region: Immunoglobulin"  
 /db\_xref="CDD:smart00409"

1359. .1586  
 /note="IG; Region: Immunoglobulin"  
 /db\_xref="CDD:smart00409"

FEATURES  
 source

gene

CDS

misc\_feature

misc\_feature

misc\_feature

misc\_feature

re 1632..1965  
/note="IG; Region: Immunoglobulin"  
/db\_xref="CDD:smart00409"  
re 1917..2072  
/note="IG; Region: Immunoglobulin"  
/db\_xref="CDD:smart00409"

56.4%; Score 43.4; DB 9; Length 2888;  
milarity 72.7%; Pred. No. 2.8e-05;  
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

3ACAAATCACAGTCTCTGGGAGGAGCATCTCTGATGCTCAGTCACTCACTCAAAAC 60  
3ACAAATCACAGTCTCTGGGAGGAGGCTCCCAAGCCCTCCATCTCCAGCAAACTCCAAAC 1629  
CTGGAGGACAAGGAT 77  
CTGGAGGACAAGGAT 1646

530 2928 bp RNA linear PAT 29-SEP-1997  
A encoding human carcinoembryonic antigen.

530  
530.1 GI:2169883  
1988177794-A/1.  
sapiens (human)  
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
nalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
(bases 1 to 2928)  
awa, S. and Nakazato, H.  
AN CARCINOEMBRYONIC ANTIGEN  
nt: JP 1988177794-A 1 21-JUL-1988;  
TORY LTD

Human  
JP 1988177794-A/1  
21-JUL-1988  
14-JAN-1987 JP 1987006851  
OIKAWA SHINZO, NAKAZATO HIROSHI  
N15/00.C07H21/04.C07K13/00.C12P21/02.G01N33/574//G01N33/577, PC  
12P21/02,  
C12K11/19;  
strandedness: Double;  
topology: Linear;  
hypothetical: No;  
anti-sense: No;  
\*source: tissue\_type=Colon tumor;  
\*source: clone=pCEA 80-11;  
Key Location/Qualifiers

mRNA 1..2929  
CDS 1..2097  
peptide 1..60 /product='human carcinoembryonic antigen' FT  
cacinobryonic antigen' FT  
mat\_peptide 61..2094 /product='signal peptide of human  
carcinoembryonic antigen' FT  
3'UTR 2099..2929 /product='mature peptide of human  
Location/Qualifiers

1..2928  
/organism="Homo sapiens"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:9606"  
56.4%; Score 43.4; DB 6; Length 2928;  
milarity 72.7%; Pred. No. 2.8e-05;

Matches 56; Conservative 0; Mismatches 21; Indels 0;  
QY 1 AGACAATCACAGTCTCTGGGAGGAGCATCTCTGATGCTCAGTCACTCAAAAC  
Db 1463 AGACAATCACAGTCTCTGGGAGGAGGCTCCCAAGCCCTCCATCTCCAGCAAAAC  
QY 61 CCGTGAGGACAAGGAT 77  
Db 1523 CCGTGAGGACAAGGAT 1539

RESULT 32  
E13123  
LOCUS 2928 bp DNA linear PAT  
DEFINITION Human gene for carcinoembryonic antigen (CEA), partial c  
ACCESSION E13123  
VERSION E13123.1 GI:3251935  
KEYWORDS JP 1997132599-A/1.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2928)  
AUTHORS Oikawa, S. and Nakazato, H.  
TITLE CARCINOEMBRYONIC ANTIGEN PROTEIN AND ITS PRODUCTION  
JOURNAL Patent: JP 1997132599-A 1 20-MAY-1997;  
SUNTORY LTD

COMMENT OS Homo sapiens (human)  
PN JP 1997132599-A/1  
PD 20-MAY-1997  
PF 14-JAN-1987 JP 1996221770  
PI OIKAWA SHINZO, NAKAZATO HIROSHI  
PC C07K14/82,C12N15/09,C12P21/02//G01N33/574,G01N33/5  
strandedness: Double;  
CC topology: Linear;  
PH Key Location/Qualifiers

FT source 1..2928  
FT /organism="Homo sapiens"  
FT /tissue\_type="colon cancer"  
FT CDS <1..2097  
FT /product='carcinoembryonic antigen'  
FT sig\_peptide <1..30  
FT mat\_peptide 91..2094  
FT /product='carcinoembryonic antigen'  
FT Location/Qualifiers

FEATURES  
source  
1..2928  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 56.4%; Score 43.4; DB 6; Length 2928;  
Best Local Similarity 72.7%; Pred. No. 2.8e-05;  
Matches 56; Conservative 0; Mismatches 21; Indels 0;

QY 1 AGACAATCACAGTCTCTGGGAGGAGCATCTCTGATGCTCAGTCACTCAAAAC  
Db 1463 AGACAATCACAGTCTCTGGGAGGAGGCTCCCAAGCCCTCCATCTCCAGCAAAAC  
QY 61 CCGTGAGGACAAGGAT 77  
Db 1523 CCGTGAGGACAAGGAT 1539

RESULT 33  
HUMCEA  
LOCUS 2929 bp mRNA linear PRI  
DEFINITION Human carcinoembryonic antigen mRNA.  
ACCESSION M15042  
VERSION M15042.1 GI:180198  
KEYWORDS Alu repeat; antigen; carcinoembryonic antigen; repeat r  
SOURCE Homo sapiens (human)

ho sapiens  
 arycota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 mallia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 (bases 1 to 2929)  
 awa,S., Nakazato,H. and Kosaki,G.  
 mary structure of human carcinoembryonic antigen (CEA) deduced  
 m cDNA sequence  
 chem. Biophys. Res. Commun. 142 (2), 511-518 (1987).  
 2844  
 .4146  
 ginal source text: Human malignant colon tissue, cDNA to mRNA,  
 nes pCEA[55-2,80-11].  
 an copy of sequence [1] kindly provided by S.Oikawa  
 (-MAR-1987).  
 Location/Qualifiers  
 1..2929  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /map="19q13.2"  
 1..2929  
 /gene="CEA"  
 <1..2876  
 /gene="CEA"  
 /product="CEA mRNA"  
 <1..2099  
 /gene="CEA"  
 /note="carcinoembryonic antigen precursor"  
 /codon\_start=2  
 /protein\_id="AAA51963.1"  
 /db\_xref="GI:180199"  
 /db\_xref="GDB:G00-119-054"  
 /translation="SAPPHRWCIQWRLLLTASLLTFWNPPTAKLTISTPFWAEG  
 KEVLVHNHUPQHLFGYSWYKGERVDGNRIIGYVIGTQATPGAYSGREIYPNAS  
 LLQNIQNDTFTLVIKISDLVNEEATQFRVYELPKPSISNNKSPVEDKDAVA  
 FTEPEQDQATYLVWNNQSLPVSPRLQSGNRLTLENVRNDTASYKCEQNPVS  
 ARRSQVILAVLXGPDAPITSPNTSVRSGENLNSCHAAENPPAQYSWVNGTFQOS  
 TQELFBNITVNSGSYTCQAHSDDTGLNRTVTITVYAEPPKPFITSNNSPVEDE  
 DAVALCEPELQNTIYLVWNNQSLPVSPRLQSGNRLTLENVRNDTASYKCEQNPVS  
 NELSDHSDFVLNVLDGPDPTISPSYTYRFGVNLSTLSCHAAENPPAQYSWVNG  
 IQQHTQLFISNTERKNSGLYTCQAANSAGHRTVTITVSAELPKPSISNNKSP  
 VEDKDAVAFTEPEAQNTYLVWNNQSLPVSPRLQSGNRLTLENVRNDTASYK  
 CGISVSNRSPVTLVDLYGPDPTIISPDSSYLSGANLNSCHASNPSPQYSWR  
 INGIPQOHTQVLFIAKITPNNGTYACFVSNLATGRNNSIVKSTITSASSTSPGNSG  
 ATVGIMTGLVGLVLI"  
 <1..91  
 /gene="CEA"  
 /note="carcinoembryonic antigen signal peptide"  
 92..2095  
 /gene="CEA"  
 /product="carcinoembryonic antigen"  
 ion 2215..2462  
 /note="truncated Alu repeat"  
 bp upstream of PvuII site.  
 56.4%; Score 43.4; DB 9; Length 2929;  
 ilarity 72.7%; Pred. No. 2.8e-05;  
 Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
 GACAATCACAGTCTCTGCGGAGCATCATCCTCTGCATGTCAGGTCAATACTCCAAAC 60  
 GACAATCACAGTCTCTGCGGAGTGCCTCAGGCTCCATCTCCAGCAACAATCCAAAC 1523  
 CCGTGGAGGACAAGGAT 77  
 CCGTGGAGGACAAGGAT 1540

VERSION AX332574.1 GI:18123208  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho  
 1  
 REFERENCE  
 AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress;  
 Horrigan,S., Soppet,D.R. and Weaver,Z.  
 TITLE Cancer gene determination and therapeutic screening usi  
 gene sets  
 JOURNAL Patent: WO 0194629-A 3093 13-DEC-2001;  
 Avalon Pharmaceuticals (US)  
 FEATURES  
 Location/Qualifiers  
 1..2974  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 56.4%; Score 43.4; DB 6; Length 2974;  
 Best Local Similarity 72.7%; Pred. No. 2.8e-05;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0;  
 QY 1 AGACAATCACAGTCTCTGCGGAGCATCATCCTCTGCATGTCAGGTCAATAAC  
 Db 1589 AGACAATCACAGTCTCTGCGGAGTGCCTCAGGCTCCATCTCCAGCAACAAC  
 QY 61 CCGTGGAGGACAAGGAT 77  
 Db 1649 CCGTGGAGGACAAGGAT 1665  
 RESULT 35  
 AX409670 2974 bp DNA linear PAT  
 LOCUS  
 DEFINITION Sequence 2317 from Patent WO0229103.  
 ACCESSION AX409670  
 VERSION AX409670.1 GI:21442375  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho  
 1  
 REFERENCE  
 AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G  
 TITLE Gene expression profiles in liver cancer  
 JOURNAL Patent: WO 0229103-A 2317 11-APR-2002;  
 GENE LOGIC INC (US)  
 FEATURES  
 Location/Qualifiers  
 1..2974  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 /note="EMBL/GenBank Accession No. M29540"  
 ORIGIN  
 Query Match 56.4%; Score 43.4; DB 6; Length 2974;  
 Best Local Similarity 72.7%; Pred. No. 2.8e-05;  
 Matches 56; Conservative 0; Mismatches 21; Indels 0;  
 QY 1 AGACAATCACAGTCTCTGCGGAGCATCATCCTCTGCATGTCAGGTCAATAAC  
 Db 1589 AGACAATCACAGTCTCTGCGGAGTGCCTCAGGCTCCATCTCCAGCAACAAC  
 QY 61 CCGTGGAGGACAAGGAT 77  
 Db 1649 CCGTGGAGGACAAGGAT 1665  
 RESULT 36  
 AX658324 2974 bp DNA linear PAT  
 LOCUS AX658324



quence 240 from Patent WO03000928.  
58324

58324.1 GI:29160809

to sapiens (human)

to sapiens

aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

lsen, H.S., Pedersen, N., Mortensen, S., Sorensen, S.B.,

ersen, M.W. and Elsnier, H.I.

hods for identification of cancer cell surface molecules and  
cer specific promoters, and therapeutic uses thereof

ent: WO 03000928-A 240 03-JAN-2003;

n Medical A/S (DK)

Location/Qualifiers

1. .2974

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

56.4%; Score 43.4; DB 6; Length 2974;

ilarity 72.7%; Pred. No. 2.8e-05;

Conservative 0; Mismatches 21; Indels 0; Gaps 0;

3ACAATCACAGTCTCTGCGAAGCATCATCTCTGCATGTCAGGTCATAACTCCAAAC 60

3ACAATCACAGTCTCTGCGAGCTGCCAAGCCTCCATCTCCAGCAACTCCAAAC 1648

3GTGAGGACAAGGAT 77

|||||

3GTGAGGACAAGGAT 1665

|||||

77147 2974 bp DNA linear PAT 27-MAR-2003

ence 340 from Patent WO02103028.

77147

77147.1 GI:29334600

to sapiens (human)

to sapiens

aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

anova, A.V., Yankovsky, N.K., Kozlov, A.P., Lobashev, A.V. and

ovskaya, L.I.

ilico screening for phenotype-associated expressed sequences

ent: WO 02103028-A 340 27-DEC-2002;

Medical Center (RU)

Location/Qualifiers

1. .2974

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

56.4%; Score 43.4; DB 6; Length 2974;

ilarity 72.7%; Pred. No. 2.8e-05;

Conservative 0; Mismatches 21; Indels 0; Gaps 0;

3ACAATCACAGTCTCTGCGAAGCATCATCTCTGCATGTCAGGTCATAACTCCAAAC 60

3ACAATCACAGTCTCTGCGAGCTGCCAAGCCTCCATCTCCAGCAACTCCAAAC 1648

|||||

3GTGAGGACAAGGAT 77

|||||

3GTGAGGACAAGGAT 1665

AX805532

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

sig\_peptide

mat\_peptide

sig\_peptide

ORIGIN

Query Match

Best Local Similarity

Matches

QY

Db

QY

Db

RESULT 39

HUMCEAF

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AX805532

Sequence 1 from Patent WO03059379.

AX805532

AX805532.1 GI:38522494

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eul

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho

Klynsner, S. and Voldborg, B.

Immunogenic CEA

Patent: WO 03059379-A 1 24-JUL-2003;

Pharmexa A/S (DK)

Location/Qualifiers

1. .2974

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

115. .2223

/codon\_start=1

/product="CEA"

/protein\_id="CAE75559.1"

/db\_xref="GI:38522495"

/translation="MESPAPPHRCIPQWRLITLTLTASLLTFWNPPT

VAGKEVLLVHNLPHLFGYSWYKGERVDGNRIIIGYVIGTQAI

PNASLLIQNIQNDGTGYFLHWIKSLDNEEATGPRVYVPELPKPS

DAVFTCEPOTDQTYLWVNNQSLPSPRLQSLNGNRILLTFNVI

NFVSARRSDSVILNVLYGDAFTISPLNTSYRGENLNLSCHAASN

FOQTQELFIPNITVNNSSYTCQAHNSDTGLNRTVTVTITVYAEF

VEDEDAVALTCEPEIQTNTYLLWVNNQSLPSPRLQSLNDNRTLT

CGIQNELSDHSDPVLNVLYGDDPTISPSYTYRPGVNLISLCH

IDGNIQHTQELFISNTEKNSGLYTCQAHNSDTGLNRTVTVTITV

NSKPVEDKDAVFTCEPEAQNTYLLWVNNQSLPSPRLQSLNGNR

RAYVCGIONSVSANRSDPVLNVLYGDDPTISPSYTYRPGVNLISL

YSWRINGIPQHTQELFISNTEKNSGLYTCQAHNSDTGLNRTVTITV

LSAGATVGMIGVLGVGLI"

115. .216

217. .2157

/product="unnamed"

2158. .2220

Score 43.4; DB 6; Length 2974;

Pred. No. 2.8e-05;

Mismatches 21; Indels 0;

1 AGACAATCACAGTCTCTGCGAAGCATCATCTCTGTCATGTCAGGTCATAAC

1589 AGACAATCACAGTCTCTGCGAGCTGCCAAGCCTCCATCTCCAGCAACAC

61 CCGTGGAGGACAAGGAT 77

1649 CCGTGGAGGACAAGGAT 1665

|||||

2974 bp mRNA linear PRI

Human carcinoembryonic antigen mRNA (CEA), complete cds.

M29540

GI:190222

carcinoembryonic antigen.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hom

1 (Bases 1 to 2974)

Barnett, T., Goebel, S.J., Notthdurft, M.A. and Elting, J.J.

Carcinoembryonic antigen family: characterization of cDN

for NCA and CEA and suggestion of nonrandom sequence var

their conserved loop-domains

**omics\_3** [1], 59-66 (1988)  
**22014**  
**0478**

```

source text: Human carcinoma cDNA to mRNA, clone pc15LV7.
Location/Qualifiers
    1..2974
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
            /map="19q13.2"
    1..2974
        /gene="CEA"
        .115..2223
        /note="CEA"
        /name="carcinoembryonic antigen"
        /codon_start=1
        /protein_id="AAA51967.1"
        /db_xref="GI:180223"
        /gb_xref="GDB:M00-119-054"
        /translation="MESFSPAPHRMCIPQWRLLLTASLTLFWNPPTTKALTSTPFPKVAEGKEVILLVHNLPQHFLFGYSWKGERVDGNRQIIGYVIQTQQATPGPAAGSKGRILLYPNALSLTIIONDTGFYLHVIKSDLVNEEATQOFRVYPPELKPSISNNKSPPVEDKDVAFTCEPETQDATALWVNNSQLPVSRLOLSNGRTLILFNTRNDTASYKCEYTNPVSARSDSVLINVLVGPDAPTICPLNSTRSGENMLSCHAASNPPAQSVSWFNVTFFQOSTGELEFINIVTNNSGSCTCAHQNSDGLNTLTVTITTYAAPPKPFITSNNSNFVEDEDVALTCETEIQNITYLMWNQSPLVPRIQLSNDNRITLLSVTRNDVGPIECGHNELSDHSDFVINLVGPDDPTISPSTYYVRGVNLRLSCHAAASNPPAOYSWLIDGNI QOHTQELEFSINTTEKNGLTVCOAKNSASHRTVKTTITVSAELPKPSISSNRSXPVEDKVDAVFCCPEAEONTVVYWNGOSLPVSRLOLSNGRTLILENVTRNDARAYVCQGINSVANSRPDLVLVGGDPTPISPPDSSYLSGANILLSCHASAASPPOKYSHRINGIQPHQTQVLFIKPNTNNGTGTAFCVENLATGRNNSIVKTSITVASGTSPGLSAGATTGMINGVLGVALL"
```

56.4%; Score 43.4; DB 9; Length 2974; Similarity 72.7%; Pred. No. 2.8e+05; Conservative 0; Mismatches 21; Indels 0; Gaps 0;	AACATCACAGTCTCTCGGAGCATCATCCTCTGCATGTGTACAGTTCATAACTCAAAC TACAATCACAGTCTCTGGGAGTGTGCCAACCTTCATCTCCAGCACACACTCAAAC GTGGAGGACAGGAT 77 GTGGAGGACAGGAT 1665
---	--

Nucleotide sequence:  
 NTSE 3036 bp DNA linear PRI 08-AUG-1995  
 un carcinobryonic antigen gene, complete cds.

Codon usage:  
 GC3.1 GI:178676  
 nioembryonic antigen.  
 ) sapiens (human)  
 ) sapiens  
 yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 alia; Buthera; Primates; Catarrhini; Hominidae; Homo.  
 bases 1 to 3036)  
 chemin,N.  
 blished (1987)  
 .bases 1 to 2541)  
 ichemin.N., Benchimol,S., Cournoyer,D., Fuks,A. and  
 mers,C.P.  
 ation and characterization of full-length functional cdNA  
 es for human carcinoembryonic antigen  
 Cell. Biol. 7 (9), 3221-3230 (1987)  
 18876  
 312  
 inal source text: Human colonic adenocarcinoma cell line Lsl80,  
 t to mRNA, and DNA.  
 t entry and computer-readable sequence for [2],[1] kindly  
 ided by N.Beauchemin, 23-NOV-1987.